Universidad Autónoma de Madrid

Departamento de Biología Molecular

Importance of post-translational modifications in stability and traffic of CAH1 and identification of chloroplast N-glycoproteins in *Arabidopsis thaliana*

Importancia de las modificaciones post-traduccionales en la estabilidad y tráfico de CAH1 e identificación de N-glicoproteínas en el cloroplasto de *Arabidopsis thaliana*

Tesis doctoral
Doctoral thesis

Andrea Martínez Bernardini

Madrid, 2014

Departamento de Biología Molecular
Facultad de Ciencias
Universidad Autónoma de Madrid

Importance of post-translational modifications in stability and traffic of CAH1 and identification of chloroplast N-glycoproteins in *Arabidopsis thaliana*

Importancia de las modificaciones post-traduccionales en la estabilidad y tráfico de CAH1 e identificación de N-glicoproteínas en el cloroplasto de *Arabidopsis thaliana*

Doctoranda:
Andrea Martínez Bernardini, licenciada en Ciencias Ambientales por la Universidad Autónoma de Madrid

Codirectoras:
Dra. Cristina Ortega Villasante. Dpto. de Biología de la Universidad Autónoma de Madrid

Dra. Flor Martínez Diez. Dpto. de Biología de la Universidad Autónoma de Madrid

Tesis realizada en el Departamento de Biología de la Universidad Autónoma de Madrid

Esta tesis ha sido realizada por la licenciada Andrea Martínez Bernardini, bajo la dirección de Cristina Ortega Villasante y Flor Martínez Diez en el Departamento de Biología de la Universidad Autónoma de Madrid

Fdo: Andrea Martínez Bernardini

En Madrid, a 13 de Mayo de 2014

Las codirectoras:

Cristina Ortega Villasante
Dpto. Biología
Universidad Autónoma de Madrid

Flor Martínez Diez
Dpto. Biología
Universidad Autónoma de Madrid

A Arsenio
A mi familia

Este trabajo ha sido financiado por:

- Proyecto BIO2009-11340, titulado "El cloroplasto de plantas como factoría para la producción de glicoproteínas recombinantes. Caraterización de una vía nativa de transporte de glicoproteínas al cloroplasto a través de endomembranas" (MINECO, Ministerio de Economia y Competitividad).

- Programa de Iniciación de Estudios de Posgrado de la Universidad Autónoma de Madrid

- Programa de Movilidad de Cultura y Deportes (ECD/3628/2011) del Ministerio de Educación y Cultura (tres meses de estancia Institute of Plant Biochemistry and Biotechnology de la Universidad de Munster, Alemania)

- ProyectoVR2008:2914,(Swedish Research Council (VR), Suecia). Dos meses de estancia en el Umea Plant Science Centre (UPSC), Suecia.

Agradecimientos/ Acknowledgements:

Hay un dicho por ahí perdido, que dice que una de las cosas importantes en esta vida es dar las gracias. Afortunadamente para mí, tengo alrededor mucha gente a la que puedo estar tremendamente agradecida.
Éste ha sido un camino muy largo, y muy difícil en algunos momentos, y me resulta muy complicado encontrar las palabras para todos los sentimientos y emociones que me vienen a la mente cuando miro el trayecto recorrido.
En primer lugar, quiero dar las gracias a Cris y Flor, mis codirectoras de tesis. ¡Al final lo conseguimos! Gracias eternas por haber enfrentado la situación que nos tocó vivir con valentía y resolución, y por haberme ayudado a sacar el trabajo adelante. Para mí sois un ejemplo de fuerza, tesón, esfuerzo y trabajo admirable. Sin vosotras esto no hubiera sido posible.
Otras personas muy importantes en el trabajo diario han sido mis compañeras de laboratorio: Amaya, "tita Magix", gracias por tu apoyo en el trabajo diario. Sabes que muchas cosas las aprendí de tí. Todavía te debo una ofrenda floral, no me olvido ;). Por eso y por todos los buenos momentos que pasamos mientras trabajábamos. Adri, Carol y Elen: gracias por vuestro compañerismo, vuestro buen humor, vuestro apoyo en ciencia y en "no ciencia" por vuestra ayuda en los momentos en los que la necesité. ¡Gracias a todas! ha sido un orgullo ser parte de las "Chlamys" durante estos años, ¡así da gusto trabajar, ha sido un placer compartir "bench" con vosotras!
Y, por supuesto, no puedo dejar de dar las infinitas gracias al principal responsable de que hoy esté aquí. Gracias Arsenio, por haber creído en mí y haberme dado la oportunidad de trabajar en tu laboratorio. Ha sido un honor haber aprendido de un gran maestro lo que es ser un buen científico. Transmitías la verdadera pasión por la ciencia, conseguías que cada experimento en el laboratorio resultara a la vez divertido e interesante. Doy por hecho que parte de todo eso lo llevo incorporado en mi forma de trabajar y afrontar los retos profesionales. Quiero pensar que estás viendo esto desde alguna parte…

Tampoco puedo dejar de agradecer a todas las personas del Departamento de Biología de la Universidad Autónoma de Madrid, que me han acompañado día a día. Gracias a Javi, Eduardo Marco, Elvira, Pili, Paco y Paqui, LuisE, Marta y Rafa, Antonio, Luis, Alfonso… seguro que me dejo algún nombre… Gracias por vuestra disposición y ayuda siempre que lo he necesitado y por vuestros consejos en el trabajo diario.

I would like to thank Prof. Göran Samuelsson for kindly reception in UPSC laboratory, for the great opportunity and learning experience that was the short stay there, and all people from the Institute: In first term to Stefan, I really remember that period as really deep learning in genomics. Thank you so much for teaching me. Besides, I want to thank all nice people in the lab: Tatiana, Juande, Aurora, Nastia…

I also would like to thank Prof. Michael Hippler for giving me the chance to perform part of my experiments at IBBP Laboratory in Münster. Special thanks as well to my diary lab-mates, Stefan and Martin, for having patience and introduce me in the crazy world of Mass Spectrometry :). Also I wanna express my gratitude to Janina, for being so friendly and help me with science issues and german translations, Leo and Deni, for your extremely nice sense of humor and help in the lab; and Philippe, Till, Ana… for diary support and interest in my work.
Y fuera del laboratorio, no puedo olvidarme de Sara y Jan, y de María y Pablo, por haber hecho de esos meses una estancia realmente agradable, por el "break" de los miércoles, por llevarme a

conocer los rincones de la ciudad y sus alrededores, y por custodiarme cuando era un peligro suelto en bici :).

Thank Dra. Muriel Bardor and María Sánchez-Contreras for kindly supervision of manuscript and helpful comments.

Gracias al resto de personas que de algún modo formáis parte de de la experiencia que ha sido realizar esta tesis: MariAngi, Aitor, Maiki, Pablo Urrita (imprescindibles sus conversaciones surrealistas en la sala de becarios), Vir y Samu, Lars, Velázquez, Ramsy, Ismael y Ana, Irene (con la que he coincidido en toda clase de momentos intempestivos trabajando en el laboratorio :D), Candela, Keyla, Gere, Isidro… A Miguelón, gracias por tu infinita paciencia contestando preguntas a esta ignorante de la informática, y por tu disposición desinteresada (o a cambio de helados) siempre a echar una mano.

También quiero agradecer a toda las bellas personas que han estado a mi lado durante todos estos años:

A mis queridísimas Paula y Blanca. Vosotras tenéis gran parte de responsabilidad de que esta tesis haya salido adelante. Dicen que es difícil encontrar amistades profundas y verdaderas, y yo me siento realmente afortunada de poder decir que os tengo a vosotras. Sois un pilar en mi vida y ¡os quiero con locura!

Mi pequeña Auro, gracias por ser el ejemplo de una amistad que nunca caduca. Gracias por transmitirme tu forma especial de ver la vida y por tu actitud optimista ante todo lo que se te cruce :). Gracias también por tu inestimable ayuda profesional y consejos con los análisis estadísticos realizados en esta tesis.

A mis cocineras favoritas Elena y Lucía, a Emilio, Nacho e Irene, Elvira, Julián, David, a mis princess del Palacio Ester y Mariona (gracias por los innumerables momentos que he compartido con vosotras. Jamás olvidaré nuestra estancia en Palacio, bizcochitos míos), y a los habitantes de Villa Manceba, Luci, Javi (aunque haya desertado a Granada) y Jean Nicolas. Gracias por compartir el día a día conmigo y tenerme paciencia en mis momentos pretésicos. Por rescatarme para desconectar cuando lo he necesitado, y por apoyarme y animarme durante el largo camino que ha sido terminar experimentos y escribir de los últimos tiempos.

También quiero dar las gracias a Mari Carmen, gran profesora de secundaria y bachillerato, la persona que me contagió su amor por la biología: ¡No me olvido de mi gran emoción cuando nos contaste la cantidad de cosas que había dentro de una célula!

Por supuesto, también quiero dar las gracias a Fer, gracias por quererme, por animarme día a día y transmitirme esa tranquilidad y paz que tanto necesito. Gracias por lanzarte a la piscina y animarte a emprender un nuevo camino conmigo :)

Por último, agradecer a mi familia: en primer lugar, gracias a mis padres, Luis y Paula. Gracias por creer en mí y apoyarme en mis decisiones. Por haberme inculcado el amor por aprender, por haberme enseñado a tener espíritu crítico. Se quedan cortas las palabras, pero si he llegado hasta aquí ha sido gracias a vosotros. ¡Os quiero muchísimo!

A mi hermana y amiga, Gabi, también me quedo corta si tengo que expresar todo lo que tengo que agradecerte. Gracias por escucharme, apoyarme y estar conmigo en los buenos y malos momentos. ¡Te adoro!
Sois el otro pilar básico de mi vida, gracias infinitas. A vosotros va dedicado todo el esfuerzo y trabajo que ha supuesto la realización de esta tesis.

Gracias a toda mi familia que vive en Argentina: vuestro cariño, aprecio y ánimos me llegan a pesar de los kilómetros que nos separan.

Importancia de las modificaciones post-traduccionales en la estabilidad y tráfico de CAH1 e identificación de N-glicoproteínas en el cloroplasto de *Arabidopsis thaliana*

Este trabajo se centra en el análisis de la N-glicosilación, una modificación post-traduccional muy común en proteínas. En concreto, se estudia el papel de la N-glicosilación en la estabilidad, tráfico y funcionalidad de proteínas cloroplásticas de *Arabidopsis thaliana*. Para ello, se realizan análisis en versiones mutadas de la glicoproteína usada como modelo CAH1, una carbónico anhidrasa codificada en el núcleo de la célula, que se transporta al cloroplasto a través del sistema de endomembranas, siguiendo una ruta diferente a la usada por la mayoría de proteínas importadas por éste orgánulo. En el proceso de transporte se modifica post-traduccionalmente adquiriendo N-glicanos y, potencialmente, un puente di-sulfuro. Se analiza la influencia de dichas modificaciones en el plegamiento y transporte de CAH1, así como la influencia del extremo carboxilo terminal.

Por otro lado, el papel fisiológico de los N-glicanos complejos, procesados y madurados al paso de la N-glicoproteína por el aparato de Golgi no está muy claro en plantas. De hecho, mientras mutantes incapaces de generar N-glicanos complejos producen fenotipos patológicos en mamíferos, mutaciones análogas en plantas no parecían presentar un fenotipo alterado. Sin embargo, estudios recientes sugieren la existencia de una relación de las modificaciones en glicosilación que tienen lugar en el aparato de Golgi, con el desarrollo normal de la planta bajo condiciones de estrés abiótico. Para profundizar en este hallazgo, en este trabajo se han utilizado plantas carentes de ciertos residuos glucídicos en sus N-glicanos, crecidas en condiciones de estrés salino, determinando diversos parámetros fisiológicos como el crecimiento de raíces o la capacidad de sintetizar pared celular.

Por último, distintos estudios indican que la ruta de transporte al cloroplasto asociada al sistema de endomembranas no es exclusiva CAH1, si no que existen otras proteínas que pueden seguir rutas similares y/o alternativas. Por ello, se planteó la identificación de otras N-glicoproteínas que tuvieran características similares a la ya mencionada CAH1. Esto implicaría no sólo la constatación de la elevada complejidad del sistema de transporte de proteínas dentro de la célula, sino la posibilidad de implementar o adaptar estas rutas de transporte a la producción de proteínas recombinantes de interés farmacéutico y/o industrial. Para ello se procedió a realizar estudios proteómicos empleando como herramienta una serie de mutantes con una ruta de transporte de moléculas al cloroplasto interrumpida.

Tanto la N-glicosilación de proteínas cloroplásticas, como los sistemas que rigen el transporte de este tipo de proteínas en plantas, se encuentran aún en estadios muy tempranos de investigación, siendo necesaria una mayor profundización tanto en la caracterización de las funciones de éstos residuos de azúcares en las estructuras proteicas, como en el tipo de sistema y elementos implicados en su transporte.

Importance of post-translational modifications in stability and traffic of CAH1 and identification of chloroplast N-glycoproteins in *Arabidopsis thaliana*

The aim of this work is to analyze the role of N-glycosylation of proteins, a common post-translational modification (PTM), on the stability, traffic and function of plastid proteins from *Arabidopsis thaliana*. With that aim, it was used as model CAH1, a nuclear encoded α carbonic anhydrase that is transported into the chloroplast via the endomembrane system, a pathway different to the canonical one used for most of the proteins targeted to this organelle. During its transport process to the organelle, CAH1 acquires N-glycans and, potentially, a disulphide bond. The role of these PTMs on folding and transport of CAH1, as well as the influence of C-terminus are analyzed.

On the other hand, physiological function of complex N-glycans, those resulting from the N-glycan processing and maturation during the pass of the N-glycoprotein through the Golgi apparatus, are not very well understood in plants. In fact, whereas mutants unable to synthesize complex N-glycans produced pathological phenotypes in mammals, in plants seems to provoke no apparently altered phenotype. However, recent studies have suggested a relationship between N-glycan modifications and normal plant development under abiotic stress conditions. To further explore this finding, we have used plants defective in specific complex N-glycan residues under salt stress conditions, and determined several physiological parameters, as root growth or cell wall formation capacity.

Additionally, several studies indicate that the route of chloroplast import of proteins associated to the endomembrane system is not exclusive of CAH1, and that there may be other proteins targeted to the organelle using similar or alternative pathways. Therefore, the identification of other N-glycoproteins sharing some characteristics than the mentioned CAH1 has risen. This finding not only would confirm the high complexity of the transport system of proteins within the cell, but it would allow adapting these routes for the production of recombinant proteins with pharmaceutical and/or industrial interest. For that, we designed a proteomic strategy, based on the use of mutants defective in a route of transport of molecules to the chloroplast.

The knowledge of both, N-glycosylation of plastid proteins and transport process of these proteins in plant cells, are still in very preliminary stages. Further research is needed to characterize the function of N-glycan residues in plants as well as the process and the elements involved on their transport to the organelle.

I.INDEX

II. ABBREVIATIONS

α-ManI:α-Mannosidase I
α-ManII:α-Mannosidase II
α(1,3)FT/FucT: α(1,3)-fucosyltransferase
α(1,4)FT: α(1,4)-fucosyltransferase
α(1,6)FT: α(1,6)-fucosyltransferase
αMM:α-methyl D-mannopyranoside
β(1,2)XylT: β(1,2)-xylosyltransferase
β(1,3)GalT: β(1,3)-galactosyltransferase
2-D DIGE/DIGE: 2-D Fluorescence Difference Gel Electrophoresis
aa: Amino acids
ADP-RF: ADP rybosylation factor
Agrobacterium tumefaciens: *Agrobacterium*
ALG: Asparagine-linked-glycosylation
Arabidopsis thaliana*:Arabidopsis*
Asn: Asparagine
BFA: Brefeldin A
BiP: Binding Immunoglobulin Protein
BSA: Bovine serum albumin
CAH1/αCAH1: *Arabidopsis thaliana*α-carbonic anhydrase 1
CAZy: Carbohydrate Active enzyme
cDNA: Coding DNA
CGD: Congenital Glycosylation Disorders
cgl1:complex glycan less mutant
CHAPS: 3-[(3 cholamidopropyl) dimethylammonio]-1-propanesulfonate
Chl: Chloroplast
Con A: Concanavalin A
CPY: Vacuolar protease carboxypeptidase Y
DEPC: Diethyl pyrocarbonate
dH_2O: Distilled water
DMSO: Dimethyl Sulfoxide
DNA: Deoxyribonucleic acid
cDNA: Complementary DNA
dNTPs: Deoxynucleotides triphosphate
DTT: Dithiothreitol
***E.coli*:** Escherichia coli
EDTA: Ethylenediaminetetraacetic acid
EGTA: Ethylene glycol tetraacetic acid
Endo H: Endoglycosidase H
ER: Endoplasmic reticulum
ERES: ER export sites
EtOH: Ethanol
FASP: Filter aided sample preparation
FT: Fucosyltransferase
For: Forward (primer)
Gal: Galactose
GFP: Green Fluorescence Protein
Glc: Glucose
GlcNAc: N-acetylglucosamine
GNTI:N-acetylglucosaminyltransferase I
GNTII:N acetylglucosaminyltransferase II
$H_2{}^{18}O$:^{18}O-labelled water
HA: Hemagglutinin
HB: Homogenization buffer
HEPES: 4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid
hgl:hybrid glycosylation mutant
HMGP: High mannose glycan protein
HPR: Horseradish peroxidase
IAA: Iodoacetamide
IgE: Immunoglobuline E
IP: Immunoprecipitation
KAc: Potassium Acetate
Kb: Kilo base

kDa: Kilo Dalton

KO: Knockout

KOR1: Endo β (1,4) glucanase

Man: Mannose

LB: Luria-Bertani medium

LC-MS or HPLC-MS: Liquid chromatography – mass spectrometry

MALDI-TOF MS: Matrix-Assisted Laser Desorption Ionisation-Time Of Flight mass spectra

MAM: Mitochondrion associated membrane

MES: 2-(N-morpholino)ethanesulfonic acid

MIMS: Membrane-inlet mass spectrometry

mM: Mili molar

MS: Mass spectrometry

MS medium: Murashige and Skoog medium

MWCO: molecular weight cutoff

Nicotiana benthamiana: *Nicotiana*

MOPS:3-(N-morpholino)propanesulfonic acid

NAA: Naphtalene acetic acid

NCBI: National Center for Biotechnology Information

NPP1: Nucleotide Pyrophosphatase/ Phosphodiesterase

OD: Optical density

OST: Oligosaccharyltransferase

pb: Base pairs

PBS: Phosphate buffered saline

PCR: Polymerase chain reaction

PEG: Polyethylene glycol

PIC: Protease Inhibitor Cocktail

PIPES: piperazine-N,N′-bis(2-ethanesulfonic acid)

PLAM: Plastid associated memebrane

pmol: Picomol

PNGasa F: Peptide-N-glycosidase F

PNGasa A: Peptide-N-glycosidase A

POX: Peroxidase

PTMs: Post-translational modification(s)

Rev: Reverse (primer)

RB: Resuspension buffer

RNA: Ribonucleic Acid

Rpm: Revolution per minutes

RuBisCO/Rubisco: ribulose-1,5-biphosphatecarboxylase oxygenase

Ser: Serine

SCs: Spectral counts

SDS: Sodium dodecyl sulphate

SDS-PAGE: SDS- polyacrylamide gel electrophoresis

SN: supernatant

SP: Signal peptide

STR: Stroma

TAE: Tampón Tris-Acético-EDTA/ tris-acetic EDTA buffer

TBS: Tris-Buffered Saline

T-DNA: Transfer DNA

TE: Total leaves extract

TEM: Transmission Electron Microscopy

***tgd*:** tri-galactosyl diaciylglycerol

TGG1/2: Thioglucoside glucohydrolase or myrosynase 1/2

Thr: Threonine

Tm: melting temperature

TP: Transit peptide

TRIS: Tris(hydroxymethyl)aminomethane

Tun: Tunicamycin

wt: Wild type

YEB:

* Abbreviations other than those listed here are defined in their first mention inthe text.

III. TABLE INDEX

Results:

Chapter 1:

Chapter 3:

IV. FIGURE INDEX

V. INTRODUCTION

1. The chloroplast and transport of proteins to the organelle

The chloroplast is an organelle of endosymbiotic origin, derived from a cyanobacteria ancestor engulfed by eukaryotic cell around 1,5 billion years ago (Douzery and Snell, 2004). In chloroplasts, different essential processes take place, such as lipid and amino acid synthesis, photosynthetic reactions or immunity responses. Due to its endosymbiotic origin, a transfer of genetic material to host cell genome took place. For this reason, only a small percentage of plastid proteins are encoded by its own genome, and consequently, most of plastid proteins are nuclear encoded. Mechanisms that facilitate translocation of these proteins into the chloroplasts evolved, and it is generally accepted that these nuclear encoded proteins are synthesized in the cytosol as precursors, with N-terminal targeting signals, named *transit peptides* (TP). Until recently, it was acknowledged that most of these proteins, were transported through the known as *canonical* pathway (Fig.I1). Here, proteins entered in the chloroplast in an unfolded way, through the so called *Toc* and *Tic* machineries, which are coupled translocators located in the outer and inner envelope membranes of the chloroplast, respectively (Radhamony and Theg, 2006; Soll and Schleiff, 2004).

However, evidences of alternative import pathways have been reported, suggesting the possibility of some proteins of being translocated by alternative pathways, other than the *canonical* one. Last decades of research point to a more complex traffic process than previously thought (Li and Chiu, 2010). For example, the *Arabidopsis thaliana* (*Arabidopsis*) prothocholorphyllide oxidoreductase A (PORA), enters the chloroplasts through an alternative translocase, other than Toc, in the outer envelope of chloroplast (Reinbothe *et al.,* 2004), also there have been reported proteins located in the inner membrane of the chloroplast lacking TP. Some identified examples are *Pisum sativum* inner envelope protein Tic 32 (Nada and Soll, 2004), and Plastid Envelope DNA-binding protein (PEND), which possess a cleavable pre-sequence potentially involved in a relocalization in the nucleus (Terasawa and Sato, 2009), and *Arabidopsis thaliana* Chloroplast Envelope Quinone Oxidoreducatse Homologue (ceQORH) (Miras *et al.,* 2002; Rossig *et al.,* 2013). In addition, a series of proteins reaching the chloroplast after passing through the endomembrane system, including both ER and Golgi, have been reported (Villarejo *et al.,* 2005; Asatsuma *et al.,* 2005; Nanjo *et al.,* 2006). This fact therefore implies the acquisition of posttranslational modifications like N-glycosylation, before entering in the organelle. Some examples are *Arabidopsis* αCAH1 (Villarejo *et al.,* 2005), α amylase I-1(Asatsuma *et al.,* 2005; Kitajima *et al.,* 2009) and Nucleotide Pyrophosphatase/Phosphodiesterase (NPP1) (Nanjo *et al.*,2006). Closely related to

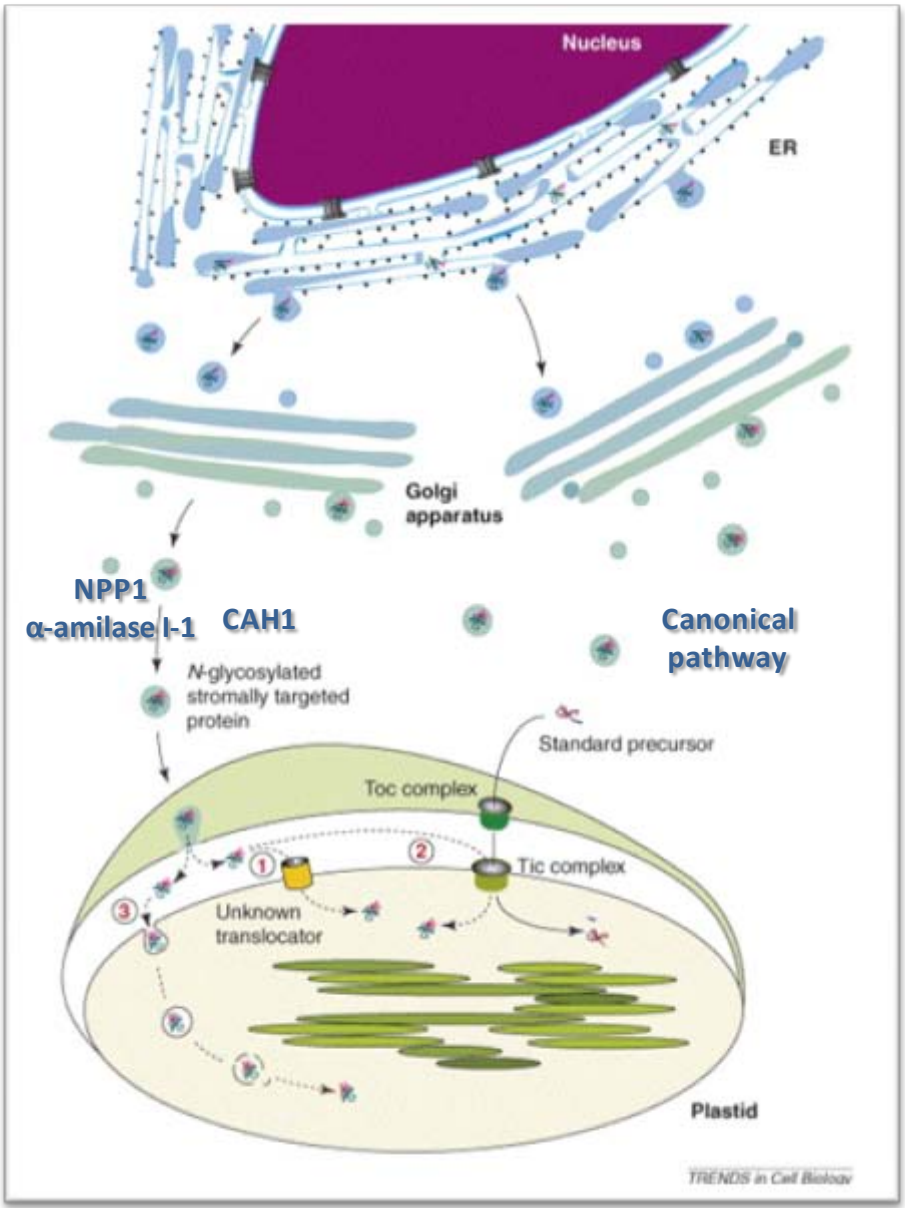

Figure I1: Canonical and non canonical pathways of transport of nuclear encoded proteins into chloroplasts. Proteins transported through indicated pathways are shown. Adapted from Radhamony and Theg, 2006.

that, these proteins possess a pre-sequence which targets them to the Endoplasmic Reticulum (ER), the *signal peptide* (SP), instead of TP, indicating their consequently obligatory pass through endomembrane system.

In fact, an increasing number of proteomic studies analyzing highly purified chloroplast samples or fractions of chloroplasts, like membranes, thylakoids or stroma, have identified a series of proteins whose structural characteristics suggest they reach the organelle trough a system other than canonical Toc/Tic. These studies performed high throughput analysis based in mass spectrometry techniques, and are consistent with the studies performed in individual cases previously named (Kleffman *et al.*, 2004; Zybalov *et al.*, 2008; Ferro *et al*, 2010; Simm *et al.*, 2013; among others).

Until now, a clear justification for the existence of these alternative pathways remains unclear. Due to chloroplast endosymbiotic origin, some studies have depicted possible scenarios considering the option that those *non canonical* import systems are reminiscences of the ancestral transport routes (Bhattacharya *et al.,* 2007; Cavalier-Smith, 2009). Interestingly, the transport of N-glycoproteins through the endomembrane system in organisms containing *complex* plastids, like those belonging to Apicomplexa or Euglenophyta phylum, among others, has been reported. Is widely accepted that *complex* plastids derived from a cyanobacteria subjected to two ancient endosymbiontic events (Gould *et al.,* 2008; Sheiner and Striepen, 2013), thus their peculiarity is that are surrounded for three or four membranes. Like in the case of *primary* plastids (like chloroplasts), a transfer of genes to the host nucleus also took place, and it has been reported that those nuclear encoded proteins reach complex plastids through the endomembrane system (Peshke *et al.,* 2013). In order to determine whether common features among species exist, the characteristics contained in the sequence of proteins subjected to this pathway have been analyzed. It was observed that they have a bi or tri partite targeting signal (Foth *et al.,* 2003), that curiously possess positive net charge as chloroplast protein TP (Waller *et al.,* 2000; Sheiner and Striepen, 2013). Some examples are the nuclear encoded proteins from *Plasmodium falciparum,* which were reported to travel through endomembrane system (Waller *et al.,* 2000; Tonkin *et al.,* 2008), and proteins from *Euglena* (Sulli and Schwartzbach, 1995). In a similar way, the occurrence of N-glycosylated proteins (protein Glx, tRNA synthetase Syn and Mutator S (MutS)) has been reported in *Phaeodactylum tricornutum,* a diatom in which proteins are transported through the endomembrane system to its complex plastids, which are surrounded by four membranes, confirming the presence of N-glycoproteins in organelles of endosymbiontic origin (Peschke *et al.,* 2013). Interestingly, there have been observed some similarities between C terminus of *Arabidopsis* αCAH1 and TP from apicomplexan N-glycoproteins, suggesting that C terminus of CAH1 could contain potential targeting information related to its pass through endomembrane system (Villarejo *et al.,* 2005). It is still unclear the mechanism used for translocation inside the organelle, although is generally accepted that Toc/Tic complex is unable to transport somehow folded proteins (Mayer, 2010). Phylogenetic data together with molecular analyses suggest that during evolution, at least in the case of *P. tricornutum*, core components of the red algal ERAD (ER-associated degradation machinery) transport system were relocalized from the symbiont's ER to the second plastid membrane to mediate preprotein translocation (Peschke *et al.,* 2013). Taken together, these observations reinforce the hypothesis that considers that this route has been maintained during evolution (Villarejo *et al.,* 2005) for proteins harboring specific posttranslational modifications needed for proper folding and/or stability (Burén *et al.,* 2011).

2. Posttranslational modifications (PTMs)

During the polypeptide synthesis process, one key step is the acquisition of posttranslational modifications (PTMs). In order to become biologically active, proteins need to acquire the proper three dimensional conformation and are subjected to some modifications, like addition of other molecular groups in specific positions (Boyle, 2005). The most outstanding PTM is N-glycosylation, widely studied in different organisms, also occurring in lipids, polysaccharides and other molecules (Varki *et al.*, 2009) and extensively treated in the present work, as well as formation of disulphide bonds. The C terminus is known to be involved in some proteins stability/folding (Wong and Ho, 2013), trafficking (Zouhar *et al.*, 2009) and activity (Pagny *et al.*, 2003). The peculiar Carboxyl terminus (C term) of CAH1, might be involved in important biological issues and contains essential information decisive for proper functioning of the protein.

For carrying out this study, two species of plants have been used as a model: On one side, *Arabidopsis thaliana*, the more widely used plant for proteomic and genetic research, due to its growth easiness, its small size, which allows easy handling, the short life cycle, and has relatively small genome, which allows a variety of experimental assays (National Science Foundation, USA, 2013). On the other hand, *Nicotiana benthamiana*, widely recognized by researchers as advantageous for transient expression assays, as it can be efficiently genetically transformed with simple methods (Goodin *et al.*, 2008).

3. The process of glycosylation

Glycosylation consists in the addition of sugar moieties to the structure of, in this case, a polypeptide, at the time or just after it is translated. It takes place in the endomembrane system of the cell, during the pass of the preprotein trough ER and Golgi apparatus. At this point, the protein can be further processed during its transport to the destination compartments, or in the destination compartment itself,

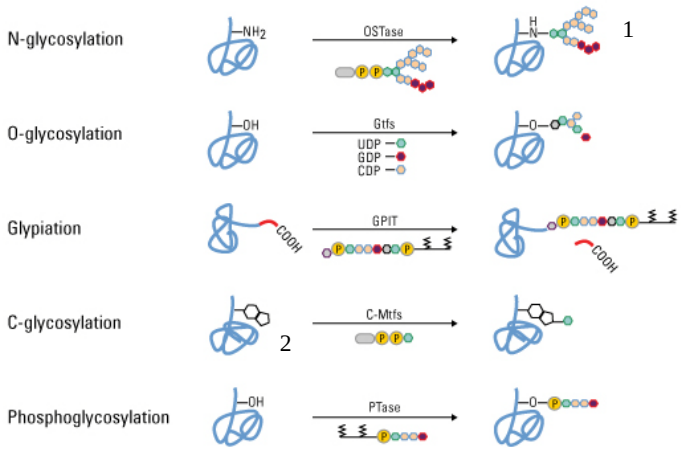

Figure I2: Types of glycosylation of proteins in cells. The different sugar structures that can be attached to proteins by different bonds are depicted. OSTase, oligosaccharide transferase; Gtfs, glycosyltransefases; U/C/GDP, nucleotide sugars; GPIT, GPI tranamidase; C-Mtfs, C-mannosyltransferase; Ptase phosphoglycosyltransferase; P, phosphorous; 1, complex N-glycan; 2, Triptophan. From Snider, 2014

like vacuoles, lysosomes and vesicles (Bardor *et al.,* 2009). Among posttranslational modifications, glycosylation is decisive for protein stability and folding. In general, glycans are involved in protein folding, cellular targeting and motility, as well as signaling for regulation of plant growth, defense and different interactions with the environment (Varki and Lowe, 2009; Price *et al.,* 2010; Larkin and Imperiali 2011). Different types of glycosylation are known: Phosphoglycosylation, C-glycosylation, glypiation and, the most outstanding, O- and N-glycosylation (Fig I2, Varki and Sharon, 2009).

In the case of O-glycosylation, an N-acetyl-galactosamine residue, followed by other sugar residues, are attached through an O-glycosidic bond to the hydroxyl group of, mainly, serine or threonine aminoacids of the primary structure of the protein. In plants, the main O-glycoproteins are hydroxyproline-rich glycoprotein (HRGPs), involved in development events and found in the cell wall, among others (Etzler and Mohnen, 2009; Gomord *et al.,* 2010). O-glycosylation is involved in folding, solubility, stability and bioactivity of proteins. Additionally it has been described to occur not only in endomembrane system, but also in cytosol or nucleus (Gomord *et al.,* 2010). On the other hand, N-glycosylation is the process where sugar residues are attached to an asparagine residue, in the consensus sequence Asn-Xaa-Ser/Thr, being "Xaa" any aminoacid but proline. The sugar residues are attached to the asparagine of the translating polypeptide in the endoplasmic reticulum, and the N-glycan is modified during the pass of the protein trough ER and Golgi apparatus and, in some cases, further modified in vacuoles, lysosomes or vesicles.

3.1 N-glycosylation in plants and other organisms

N-glycosylation occurs throughout the different kingdoms in nature, that is, *Bacteria, Archaea* and *Eukarya*. The variability that exists confers specific characteristics and particularities to the proteins subjected to these PTM (Larkin and Imperiali, 2011).

Synthesis of N-glycoproteins (see Fig I3 for details) starts in ER, where the polypeptide acquires, co-translationally, a standard glycan precursor in its Asn residues within the previously mentioned consensus sequence (Gomord *et al.,* 2010; Fig I3a). This precursor is further modified during the transport of the protein through the ER and Golgi apparatus by the sequential action of ER resident glycosidases (Fig I3b) and Golgi glycosidases and glycosiltransferases (Fig I3c). When processed in the ER, terminal glucoses are trimmered from the glycan precursor, giving rise to a high mannose type N-glycan (Fig. I3d). This high mannose precursor is further processed when, subsequently, protein pass through Golgi system. Here, some mannose residues are trimmered, and GlcNAc in α(1,3) position is added, by the action of N-acetylglucosaminyltransferase I, GnTI (Fig I3e). GnTI is the key enzyme involved in complex glycan synthesis. The process, until this step, is highly conserved among the different eukaryotic organisms (Kelleher and Gilmore, 2006; Mohorko *et al.,* 2011 ; Zielinska *et al.,* 2012).

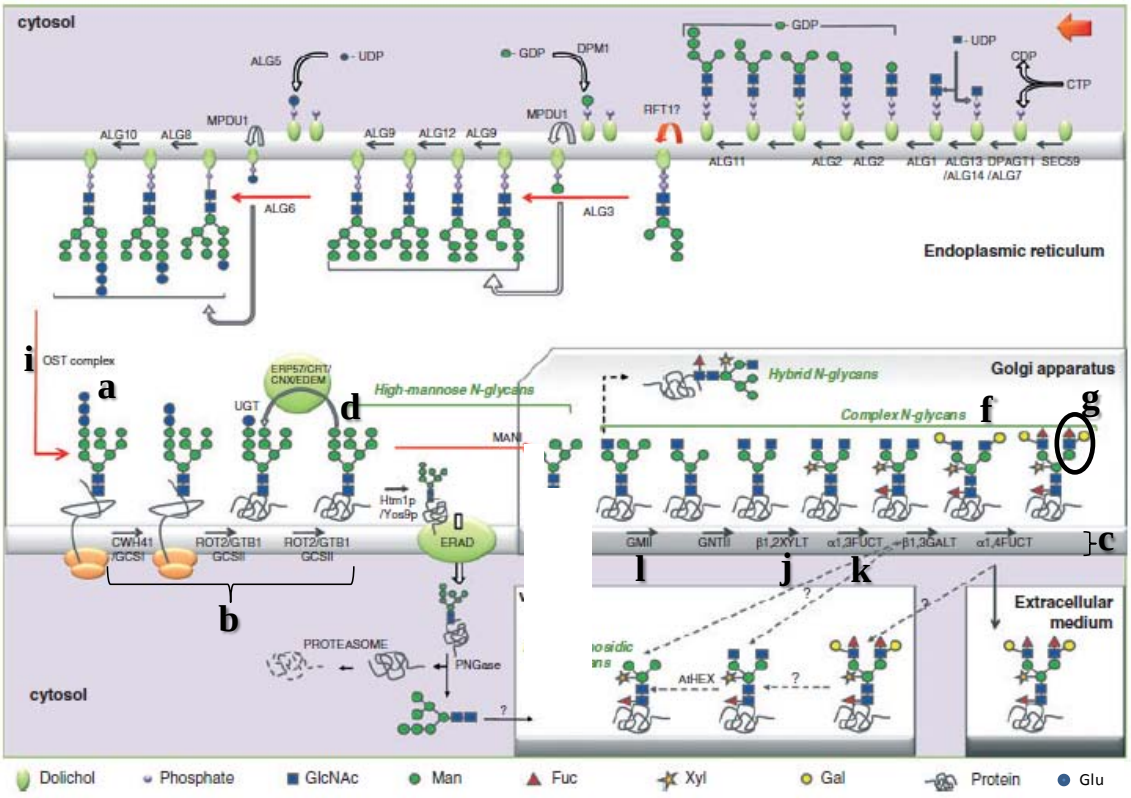

Figure I3: The process of N-glycosylation in plants. a) Linking of standard glycan precursor to the nascent protein, b) Glycosydases removing glucose moieties from the standard precursor, c) Glycosyltransferases processing N-glycan into hybrid or complex type, d) High mannose N- glycans, e) GnTI enzyme, f) Complex N-glycans , g) Lewis[a] structure, h) Paucimannosidic N-glycans, i) OST complex, j) β(1,2)XylT enzyme, k) core α(1,3)FucT enzymes, l) α mannII enzyme. Adapted from Gomord *et al.,* 2010

However, process differs in subsequent final steps among different organisms, where different sugar residues are added to the standard N-glycan precursor. As a result, processed complex N-glycan vary depending on the number and type of residues that is finally harboring (Fig I3f and Fig. I4). Interestingly, plants and insects share some similarities, when compared to vertebrates, as both contain α(1,3)fucose linked to the proximal GlcNAc in the core of chitobiose in the structure. Additionally, plant complex N-glycans also contain β(1,2) xylose residues, which are occasionally present in some snails/invertebrates (van Kuik *et al.*, 1985) and parasitic trematodes (Faveeuw *et al.*, 2003; Bencurova *et al.*, 2004). Both core α(1,3)fucose and β(1,2)

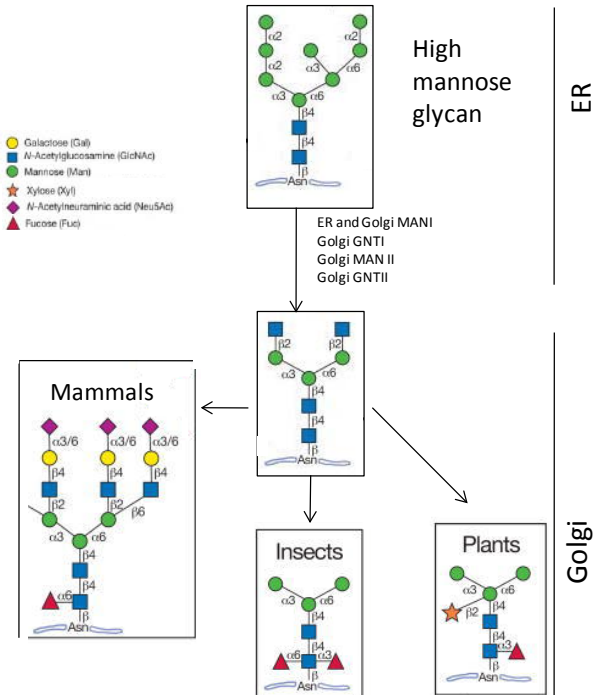

Figure I4: Complex glycan processing in different organisms. Adapted from Varki *et al., "Essentials of glycobiology"*, 2009

xylose are absent in mammals. In the case of vertebrates, complex glycans are bigger and more heterogeneous structures, decorated with sialic acid, β(1,4)galactose and α(1,6)fucose (Raikhel and Chrispeels, 2000; Wilson, 2002), which are absent in plants. In any case, N-glycoproteins can be further processed, acquiring branched structures known as Lewis structures (Lewis^a in the case of plants, Fig. I3g) (Wilson, 2001). They can also be further modified *en route* to the targeting destination, or within the final compartment itself. In plants, a common structure illustrative of this case are paucimannosidic N-glycans, which are modified in the vacuole or after processing Lewis^a branched structures (Fig. I3h).

3.1.1 Biological significance

As mentioned before, N-glycosylation is widely known as essential for proper folding and function of proteins. Indeed, early N-glycosylation, performed in ER, is highly important for both, mammals and plants, as shown by the highly conserved machinery in different organisms along evolution. Several studies in mammal systems have demonstrated that is indispensable for a series of biological processes, as mutation in key ER enzymes involved in N-glycosylation leads to serious disorders known as *Type I Congenital Glycosylation Disorders* (CGD) (Freeze, 2002; Lowe and Marth, 2003).

Specifically in plants**,** early glycosylation is known to play pivotal roles in abiotic stress tolerance, cellulose biosynthesis and cell wall formation, normal embryogenesis and other plant development events (Wilson, 2002), and also in detection of microbe-derived molecules, triggering immune responses (Saijo *et al.*, 2010). Mutations in some processing enzymes at this stage, like components of the *OST* (oligosaccharyltransferase complex sub-units), which is responsible for the transfer of the oligosaccharide precursor onto the nascent glycoprotein (Fig. I3i), can cause even embryo lethality, as demonstrated by several studies (Boisson *et al.*, 2001; Koiwa *et al.*, 2003; Lerouxel *et al.*, 2005; Saijo *et al.*, 2010).

Regarding late N-glycosylation, which takes place in Golgi system (Fig.I3), studies in animal models have shown the role of *complex* N-glycosylation in protein stability and folding, cell adhesion and recognition, proliferation and differentiation, immunity, development (Ioffe and Stanley, 1994 ; Lau *et al.*, 2007), and proper tissue formation (Ioffe *et al.*, 1996). Disruption in Golgi N-glycosylation of mammal/vertebrate glycoproteins leads to developmental disorders, at tissue or organ level, and even embryo lethality. Particularly in humans, mutations in Golgi resident glycosyltransferases lead to the so called *Type II* CGD, resulting in pathological disorders, some of them lethal (Wang, 2001; Freeze, 2002; Freeze and Aebi, 2005). In other organisms, like *Drosophila melanogaster*, functional N-glycoproteins are essential for the development and central nervous system (Sarkar *et al.*, 2006) and in *Candida albicans* these proteins are required for cell wall integrity (Bates *et al.*, 2006). However, in plants, the presence

of the cell wall seems to hamper some of the N-glycan's functions previously listed in animal cells, such as cell adhesion. Intriguingly, no clear explanation for the function of plant late N-glycan processing in Golgi apparatus has been raised. Mutations in plant Golgi resident glycosyltransferases do not produce conspicuous changes in plant phenotype, when grown in normal conditions. Up to date, neither their function nor the exact implications of glycan's different structures are completely clear. Plant mutants disrupted in late steps of the N-glycosylation process, as the *Arabidopsis* complex glycan less (*cgl1*) mutant, a GntI (Fig. I3e) knockout, or knockouts in xylosyl-or fucosyltransferases (Fig. I3j and k, respectively), like *FucTA/FucTB/XylT*, do not show any evident phenotype in normal growing conditions (von Schaewen *et al.,* 1993; Strasser *et al.,* 2004; Strasser *et al.,* 2006). Nevertheless, in spite of the apparent absence of function of plant complex N-glycans, some studies have reported that, under abiotic stress, complex N-glycosylation is required for proper functionality of proteins (Kang *et al.,* 2008), implying a potential involvement of complex N-glycans in some kind of protection to the proteins, under adverse environmental conditions. This protective role is somehow confirmed in the *Arabidopsis* mutant knockout in Golgi Man II (Fig. I3l), hybrid glycosylation *(hgl*), who has α(1,3)fucose and β(1,2)xylose residues but they are not accessible to specific antibodies or human Immunoglobuline E (IgE), producing low immunogenic response, and consequently demonstrating the involvement of these residues in allergenic responses (see section 3.1.2). On the other side, the *cgl1* mutant presented higher sensitivity to salt stress, in accordance with the previously mentioned potential involvement of glycans in abiotic stress tolerance (Kang *et al.,* 2008; Kaulfürst-Soboll *et al.,* 2011).

In this line, there are evidences reporting that salt stress triggers cell wall defects in mutants with deficient N-glycosylation, such as the mentioned *cgl1* (Kang *et al.,* 2008). The importance of N-glycosylation in enzymes involved in cellulose -a cell wall core component- biosynthesis has been demonstrated (Lukowitz *et al.,* 2001; Boisson *et al.,* 2001; Burn *et al.,* 2002; Gillmor *et al,* 2002; Kang *et al.,* 2008; Cubero *et al.,* 2009; von Schaewen *et al.,* 2008). One N-glycoenzyme involved in the synthesis of cellulose is endo β(1,4)glucanase / KORRIGAN 1 (KOR1) (Nicol *et al.,* 1998; Sato *et al.,* 2001; Endler and Persson, 2011), whose precise molecular function remains unclear, but its N-glycosylation state has been shown to influence cell wall functionality under salt stress (Kang *et al.,* 2008).

3.1.2 Immunogenicity of α(1,3)fucose and β(1,2)xylose

Plant N-glycoproteins have specific complex N-glycans structures, containing β(1,2)xylose moiety, attached to the β linked mannose residue, and core α(1,3)fucose moiety attached to the proximal GlcNAc, in the common conserved $Man_3GlcNAc_2$ core (Fig. I4). As mammals lack xylosylation and α(1,3)fucosylation, these epitopes are recognized by mammalian immune

systems as foreign elements, then triggering immunogenic responses (Wilson , 2002; Bardor *et al.,* 2009; Both *et al.,* 2011). A similar reaction is triggered by N-glycans found in insects, as they also harbor core α(1,3)fucose (Rendic *et al.,* 2007). The antigenicity of complex plant N-glycans is well documented (Foetisch *et al.,* 2003; Bencurova *et al.,* 2004). However, from a biotechnological point of view, is a handicap for production of therapeutic proteins in plants that has to be considered, for example, in treatments requiring large quantities of repeatedly administered glycoproteins (Sriraman 2004; Bardor *et al.,* 2009). See section 5 for more information about use of plants as biofactories.

4. The plant cell as a dynamic system: inter-membrane connections and proteome plasticity

It has been accepted, since several years ago, the idea of living cells as continuously changing entities, in which external and internal stimuli are in permanent feedback. Clear evidence of this fact are the signaling pathways, in some cases triggered by situations like changes in the media or presence/absence of potential interacting organisms, enabling the cell to respond in order to face the changes. Then, there should be some communication *highways* within the cell, in terms of chemical signals and/or regulation of (some) protein levels. Such communication may be quite direct and fast, and the question is how it can be achieved.

4.1 Communication between membranes

There is increasing evidence of tight contact between different parts of the cell. It has been characterized, among different organisms, a network formed by the ER that connect organelles like Golgi, mitochondria, plasma membrane, nucleus, and chloroplasts in the case of plants (Griffing, 2011). For example, direct contact between ER and mitochondrion has been demonstrated in yeast (Achleitner *et al.,* 1999). In plants, although scarcely characterized, some examples have been reported of direct contacts between ER and mitochondria or chloroplasts.

The presence of N-glycosylated proteins in the chloroplast, like those published by Chen *et al.,* 2004; Nanjo *et al.,* 2006; Asatsuma *et al.,* 2005 or Villarejo *et al.,* 2005, establish the idea of not previously considered connections between organelles and endomembrane system. However, it remains unknown the nature of these connections.

As previously mentioned, there are different experimental evidences pointing to the dynamism of membranes. This would mean they are interconnected, forming a network within the cell, with a continuous traffic of molecules from some compartments of the cell to others. In this line, changes in permeability of outer membrane of chloroplasts under some stress conditions have been observed (Krause *et al.,* 2012). Besides, as previously mentioned, there are evidences of

direct connections between membranes, as ER-Mitochondria**,** *Mitochondrial Associated Membranes* (MAMs) (Rowland and Voeltz. 2012); or ER- Chloroplast, *Plastid Associated Membranes* (PLAMs) (Kelly and Dörmann, 2004; Benning *et al.*, 2006; Andersson *et al.*, 2007; Mehrshahi *et al.*, 2013). It has been suggested that PLAMs are the passing gateway of lipid precursors assembled in the ER needed for synthesis of thylakoid membranes inside chloroplast (eukaryotic pathway), together with those lipids that are synthesized *de novo* inside the chloroplast (prokaryotic pathway) (Xu *et al.*, 2008; Benning , 2009) (Fig. I5). The contribution of these two pathways, chloroplastic/prokaryotic and eukaryotic, to thylakoid lipid biosynthesis, are nearly equal (Awai *et al.*, 2006).

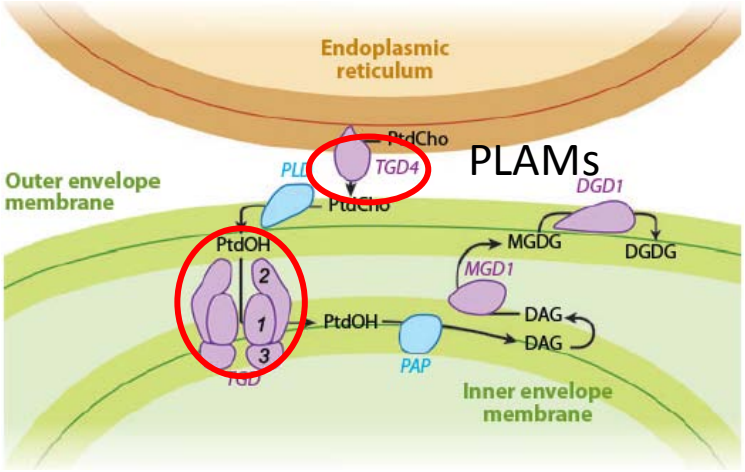

Figure I5: Proposed model for lipid precursors import to chloroplasts. Model for eukaryotic lipid precursors import in *Arabidopsis thaliana* (eukaryotic pathway): Fatty acids are *de novo* synthesized in the chloroplast and exported to the ER. There are converted into phosphatidylcholine (PtdCho), which is transported to chloroplast in a TGD4 mediated form, and subsequently transformed in phosphatidic acid (PtdOH), by the action of phospholipase D (PLD). PtdOH is imported to the chloroplast through the TGD translocon (formed by TGD1/2/3 proteins) located at the inner membrane. There, by the action of phosphatidic acid phosphatase (PAP), give rise to a diacylglycerol (DAG) moiety, which is further incorporated into monogalactosyldiacylglycerol (MGDG), by the action of glycosyltransferase MGD1, or into digalactosyldiacylglycerol (DGDG), by the action of glycosyltransferase DGD1. PLAMs, Plastid Associated Membranes. Adapted from Benning, 2009.

Transport of molecules other than lipid precursors through PLAMs still remains unexplored, but some mutants impaired in the eukaryotic pathway of lipid precursors import accumulate galactoglycerolipid trigalactosyldiacylglycerol (TGDG) in the ER, and present some altered phenotypic features when compared to the wild type, like slow growth, paler colour, oligogalactolipids accumulation, reduction in eukaryotic lipid synthesis or alteration of chloroplast structure (Benning *et al.*, 2006). These facts clearly point to the TGD translocon as an essential element for chloroplast integrity, and lead to consider the possibility of other type of molecules entering tho chloroplast through this pathway.

Additionally, it has been reported the existence of *stromules*, direct connections between chloroplasts (Kwok and Hanson, 2004) which are also tightly related to the ER (Schattat *et al.*, 2011). Furthermore, it was observed that the plant Golgi apparatus has not a fixed position inside the cell, so individual Golgi stacks are moving rapidly along actin microfilaments that lie parallel to ER tubules (Lerich *et al.*, 2012). This suggests a tight coupling between ERES (ER Export Sites) and Golgi, leading to the concept “secretory unit” (daSilva *et al.*, 2004; Lerich *et al.*, 2012). Moreover, direct tubular connections between ER and Golgi have been reported (Griffing 2011; Hawes 2012). Taking all together, it seems there is a significant intercommunication within the

cell, although the underlying mechanism is still discussed and several models have been proposed. Nevertheless, all evidences points to a very dynamic system highly intercommunicated.

4.2 Chloroplast proteome: dynamism, composition and presence of N-glycoproteins

Considering the cell as an active organized structure in permanent contact with the surrounding environment, a moderate number of studies assessing the variation in chloroplast proteome have been carried out in the last decade. They have shed light on changes in the protein pool under different internal and external conditions, like abiotic stresses such as salinity, drought or flooding, or changes in temperature (Goulas *et al.,* 2006; Hossain *et al.,* 2011). Reported variations in general proteome levels, in relation with mentioned media changes, are in agreement with the already mentioned communication between membranes.

Focusing in protein composition, without considering any interaction with the media, a series of studies revealed new insights in the composition of chloroplast proteome. For example, as mentioned before (see section 1), the presence of nuclear encoded plastid proteins possessing SP instead of TP have been reported, being many of them not described as plastid resident before (Kleffmann *et al.,* 2004; Zybalov *et al.,* 2008; Armbruster *et al.,* 2009; Ferro *et al.,* 2010). This suggests the existence of alternative targeting systems and pathways, supporting potential communication/contact with ER. Focusing in N-glycoproteins, first studies indicated a very low number of N-glycoproteins following the endomembrane system route (i.e, similar than αCAH1) (Zybalov *et al.,* 2008). Low amount of these type of N-glycoproteins was further confirmed by specific analysis of plastid N-glycoproteome (Ruiz-May *et al.,* 2012; Zielinska *et al.,* 2012). Estimated percentages of non typical plastid proteins found in the organelle varies among studies comprising the whole chloroplast, although is around 12%, in agreement with studies considering exclusively chloroplast subcompartments (Friso *et al.,* 2004; Giacomelli *et al.,* 2006; Peltier *et al.,* 2006). Among the already described plastid N-glycoproteins it has to be noticed that they seem to exhibit differences in the N-glycan composition. For example, AmyI-1 presents partial sensitivity to EndoH and is totally digested by PNGase F (Asatsuma *et al.,* 2005), in accordance to the xylose-containing carbohydrates reported by Hayashi *et al.* (1990). In the same line, CAH1 protein was reported to contain both α(1,3)fucose and β(1,2)xylose residues (Villarejo *et al.,* 2005). However, NPP1 shows complete susceptibility to Endo H and affinity to Concanavaline A (Nanjo *et al.,* 2006), meaning that it probably bears only high-mannose-type N-glycans, typical from ER. This suggest a direct ER-chloroplast transport of high-mannose-type N-glycoproteins as for lipid precursors, via PLAMs. In fact some studies have speculated with the possibility of a direct transport of proteins between ER and chloroplasts (Andersson *et al.,* 2007; Bhattacharya *et al.,* 2007; Tan *et al.,* 2011; Krause *et al.,* 2012; Gagat *et al.,* 2013).

Additionally, a potential set of proteins dual targeted both to mitochondria and chloroplasts have been identified, confirming variations in proteome composition and suggesting targeting to different organelles as more frequent than previously thought (Ferro *et al.,* 2010), and therefore highlighting the dynamism of membranes and proteome.

4.3 Dual targeting

It is known as *dual targeting* the process in which some proteins are delivered to two or more different destinies within the cell. There are increasing evidences of proteins subjected to this type of transport, not only in plants but in several eukaryotic organisms (Karniely and Pines, 2005). However, the exact amount of proteins with this special feature still remains unclear, neither are known the reasons of the existence of this mechanism. In plants, dual targeting to chloroplasts and mitochondria has been reported, as a control system of the biogenesis and function of these organelles (Carrie and Small, 2013). Other authors propose dual targeting to chloroplast and nucleus as a regulatory element as signals from the chloroplast, in order to face some environmental stress (Krause *et al.,* 2012). Regardless the primary reason, it is not surprising the potential occurrence of such a variable transport, taking into account the already mentioned evidences of dynamic, continuous and adaptable traffic and communication system within the cell. That leads to the previous concept of the cell as a wide network interconnected, continuously changing, within the organelle communications adapting to the moment needs. In *Arabidopsis*, it has been estimated that there are around 100 proteins targeted both to mitochondria and chloroplast (Carrie *et al.,* 2009), although there are also evidences of proteins being targeted to chloroplast and nucleus (Schwacke *et al.,* 2007) or the peroxisome (Sapir -Mir *et al.,* 2008). Regarding other organisms, there are also studies in the model unicellular algae *Chlamydomonas reinhartii* in which protein disulfide isomerase RB60, is targeted both to ER and chloroplasts (Levitan, 2005), or *Saccharomices cerevisae,* in which it was clearly quantified that around a third of mitochondrial proteome is also located in other compartments of the cell (Ben-Menachem *et al.,* 2011).

4.4 Location of glycosyltransferases within the Golgi apparatus: a controversial issue

Taking together the increasing evidences of dynamism between membranes, and the changes suffered by proteome according to environmental interactions, it can be expected that those changes would be reflected also at membrane level. As the present work is focused in the particular process of N-glycosylation, special attention has been paid to localization of the proteins involved in maturation of N-glycans, i.e those located along the Golgi system. Up to date, ambiguous data point to localization of glycosyltransferases as a controversial matter. There

are different hypothesis explaining the way glycosyltransferases attach to the Golgi membrane, as evidences indicate that, on one hand, they act in a fixed sequential manner and, on the other hand, that they are able to catalyze reactions in an alternative order than the one generally accepted (Lerouge *et al.,* 1998).

Some studies reported that *in vitro*, β(1,2)xylose and α (1,3)fucose residues can be added at early stages in the biosynthesis pathway, suggesting the possibility of alternative routes occurring *in vivo* (Bencúr *et al.,* 2005; Strasser *et al.,* 2006; Kajiura *et al.,* 2012). Moreover, other studies propose that fucosylation and xylosylation reactions are independent (von Schaewen *et al.,* 1993; Forth et al., unpublished), supporting the idea of these enzymes acting at any stage of the pathway.

However, it seems that sequential action of enzymes is preferred *in vivo*, due to the specific Golgi localization of the enzymes (Schoberer *et al.,* 2013). In this line, there are evidences of multiple mechanisms functioning to ensure the correct localization of glycosyltransferases. Tu and coworkers (2008), proposed that enzymes maintain their localization in the Golgi in a steady-state way, being recycled through iterative rounds of retrieval and forward vesicle-mediated transport. In addition, it has been observed that *cis* and *medial* Golgi enzymes tend to interact with each other *in vivo*, through their N terminal *c*itoplasmic *t*ransmembrane *s*tem region (CTS), giving rise to homodimers and heterodimers that are subsequently blocked in Golgi membranes. These unions vary the strength throughout the membrane, leading to a gradient-like fashion distribution of the enzymes (Tu and Banfield, 2010). In accordance to this proposed fixed position of glycosyltransferases in the membrane, and consequently sequential action, it has also been observed that core fucose addition strictly requires previous xylosylation in various plant species, including *Arabidopsis* (Kaulfürst- Soboll, *et al.,* 2011). Nevertheless, how exactly the process occurs and if observations *in vitro* can really take place in natural conditions, still remain unclear.

The particular importance of fucosyltransferases (FTs from now on) lies in that they catalyze the final, and sometimes crucial, step in the synthesis of a range of biologically active glycoconjugates. Fucosylation has been extensively studied in other organisms as it is involved in a variety of biological processes and it is known as disease bioindicator in humans/mammals systems (Both, 2010). In addition, a deep characterization would facilitate the creation of genetically modified plants, as fucosylation events could be accurately predicted, permitting both avoidance of fucose attachment in recombinant glycoproteins, as well as addition genetic modifications required to produce humanized structures for therapeutic proteins, like Lewis $^{x/y/b}$ N-glycans (Bakker *et al.,* 2001). According to CAZy database (cazy.org, Couthinho *et al.,* 2003), FTs are divided in two superfamilies, one containing α1,3/4 FTs, (GT10 family) and another including FTs activities classified in families 11, 23, 37, 65 and 68 (Both *et al.,* 2011).

Regarding the controversial localization of enzymes catalyzing last steps in N-glycosylation, and particularly core α(1,3) fucosylation, is noticeably that is the only step carried out by two

enzymes, FTs 11 and 12 (FT11 and FT12, from now on), neither their specific action nor the exact localization is still fully understood. It is not clear whether they are functionally redundant or they act in processes taking place in different Golgi stacks. Inactivation of both, FT11 and FT12 is necessary to completely abolish core activity. Strasser and coworkers (2004) reported that the activity of only one of them seems to fully complement the absence of the other. Nevertheless, previous studies performed in our laboratory as well as in that of the Prof. G. Samuelsson in the Umeå Plant Science Center (UPSC, Sweden) (Blas, unpublished; Forth *et al.,* unpublished, respectively) showed clear differences on the α1,3 fucosylation levels of N-glycoproteins targeted to the chloroplasts in *Arabidopsis* FT12 knockout *(AtFT12)* suggesting a specialization of this isoform in α(1,3)fucosylation of chloroplast proteins, while little effect was observed on the α(1,3)fucosylation pattern of other sub-cellular fractions. The accurate localization of FTs would allow to find out whether FTs are sequentially localized in Golgi, which is presently the most accepted hypothesis, or specialized Golgis, processing proteins with a specific destination exist, as other studies suggest (Forth *et al.*, unpublished).

5. Biotechnological applications: plants as biofactories

Interest in PTMs processes in plants is not only due to basic knowledge about cell functions gaining, but also for the use in potential biotechnological applications. Since N-glycosylation is present in most proteins of therapeutic and/or industrial applications, the possibility of using plants to produce proteins of pharmaceutical aptness, as vaccines or antibodies, or with industrial purposes, like food crops with reduced immunogenicity (von Schaewen and Frank, 2008), has raised. Up to date, the standardized expression systems for recombinant proteins production are based in animal/mammal models, being the most extended *Chinese Hamster Ovary* (CHO) cell lines. Additionally, alternative systems like insect cells, yeast, filamentous fungi or even derived-human cell lines and transgenic animals are being developed (Durocher and Butler, 2009), although these methods present several disadvantages. In all cases it still is necessary to engineer the pathway, since N-glycosylation is not always exactly the same as in humans. However, in the case of mammalian cell lines, their genetic manipulation is more complex. Besides, the protein products obtained present highly heterogeneous N-glycosylation, thus being low reproducible. Additionally they present high production costs and risk of contamination with human pathogens (Bosch *et al.*, 2013).

For that reason, the possibility of using photosynthetic organisms like algae (Mayfield *et al.*, 2007) or plants, instead of animal systems, as biofactories arose. Indeed, producing recombinant proteins in plant systems present economic and practical benefits. They are versatile systems easily adaptable, by genetic modifications, to produce N-glycoproteins. In addition, the possibility of a quick scale-up of production maintaining moderate costs still exists. Additionally, these

systems are absent of human pathogens (Ma *et al*., 2003; Daniell *et al.*, 2005; Bosch *et al*., 2013). Currently, plants containing genetically modified N-glycosylation pathways have been created, following different strategies. Some are devoid of immunogenic complex N-glycans by *knock out* techniques (Bardor *et al.*, 2009), or retention in ER by KDEL/HDEL peptide fusions (Sriraman *et al.*, 2004). Different recipient species are being investigated, as the moss *Physcomitrella patens* (Koprivova *et al*., 2004) or *Nicotiana* (Strasser and Stadlmann, 2008). In general, plant systems present other problems in addition of immunogenic residues, as low expression levels or low stability of the heterologous proteins. In this line, efforts are focused in improvement of the yield of the production.

For that, one of the options considered is the possible accumulation and further purification of the recombinant proteins in the chloroplast, either by stable transformation (Fernandez-San Millan *et al.*, 2008; Gao *et al.*, 2012) or by targeting to the organelle. Internal machinery of this organelle is able to process some eukaryotic proteins, although it seems to lack machinery involved in N-glycosylation. Additionally it can accumulate proteins up to 10%- 40% of total soluble protein (Daniell, 2005; Daniell *et al.*, 2009). Together with the reported non canonical routes (Asatsuma *et al.*, 2005; Villarejo *et al.*, 2005; Nanjo *et al.*, 2006), it seems possible to use chloroplasts to improve the yield and stability of N-glycoproteins of interest. Nevertheless, as previously mentioned, the addition of immunogenic α(1,3)fucose and β(1,2)xylose residues should be avoided (Strasser *et al.*, 2004). Different approaches have been developed to solve that problem, as the removal by knocking out of both α(1,3)fucosyltransferases and β(1,2)xylosiltransferase (Strasser *et al.*, 2004). On the other hand, addition of mammalian glycosyltransferases to achieve humanized pathways in different species like *Withania somnifera* or *Nicotiana* has also been achieved (Sriraman *et al.*, 2004; Jez *et al.*, 2013; Castilho and Steinkellner, 2012). Some plant models are already being used for the production of pharmaceutical proteins, and some plant-derived recombinant proteins had been approved for human healthcare with positive results: antibody CaroRx, oral and injectable vaccines (Davoodi-Semiromi *et al.*, 2009; Davoodi-Semiromi *et al.*, 2010), therapeutics like interferon alpha in *Lemna minor* or glycocerebrosidase prGCD in carrot cells (Shaaltiel *et al*., 2007).

However, it would be very useful for these kind of biotechnological applications to find a way in which proteins only harbouring high mannose type N-glycans could be directly transported to the chloroplast, since this sort of N-glycan structure is common to all eukaryotic organism and does not raise allergic immunogenic reactions in humans. Then it would only be necessary to add the human/mammal specific glycosyltransferases, like sialyltransferase to have a humanized therapeutic protein.

It is in this point where basic research and biotechnology join together, highlighting the need of a more detailed knowledge of the plant N- glycosylation process, in order to improve the strategies

that enable heterologous proteins production with a specific mammalian/human-like N-glycosylation.

VI. OBJECTIVES

As every breakthrough discovery in science, the description by Villarejo *et al.*(2005) of a novel of chloroplast protein import pathway for nuclear encoded proteins, as the α carbonic anhydrase CAH1, after its synthesis and post-translational modifications in the endomembrane system, generated a great number of new questions. The role of N-glycosylation and other posttranslational modifications, or the importance of peculiar protein sequence motifs as the C terminus on CAH1proper folding and transport, should be examined to understand how these proteins are targeted to the chloroplast. On the other hand, a general knowledge of plant specific N-glycosylation processes and function might shed light on the biological basis for the existence of this pathway.

The general aim of this work was to address some of the questions raised above. To achieve this objective, the following specific goals were established:

1. Unravel the importance of CAH1 post-translational modifications, namely N-glycosylation, potential disulphide bond formation, and the influence of the C terminus, on protein structure, trafficking and function.
    - A strategy based on the generation of point mutations in a haemmaglutinin tagged CAH1version (HC) was designed. Mutations on the five potential N-glycosylation sites, as well as on four cysteine positions were created. In order to test whether the sequence or the presence of the C terminus itself were relevant for CAH1 proper functioning, point mutations on specific lysine residues and a complete deletion of this domain were also generated.
    - The series of HC mutated versions were transiently expressed in different systems: *Arabidopsis thaliana* cell cultures and plant mesophyll protoplasts, versus *Nicotiana benthamiana* leaves, and biochemical and immunological analyses were performed.

2. Analyze the role of plant specific complex N-glycosylation from a physiological point of view, by testing abiotic stress tolerance, with special attention on fucosylation events.
    - In order to determine the contribution of plant complex N-glycans, or specifically α(1,3) fucose residues to abiotic stress tolerance, a comparative study using mutant lines altered in different steps of the N-glycosylation process treated with different concentrations of salt was performed. Plant growth and a series of parameters related to cell wall consistency were tested in *Arabidopsis thaliana* wild type Columbia 0 (wild type/wt), the α(1,3)fucosyltransferase double mutant (ΔFT), defective in the

addition of core α(1,3)fucose residues and the *complex glycan less 1* mutant (*cgl1*), which lacksplant specific complex N-glycans.

- Focusing on fucosylation, the precise subcellular localization of the two proteins that perform core α(1,3)fucosyltransferase activity in *Arabidopsis* was studied in order to determine whether they co-localize in similar or conversely, different sets of Golgi stacks. For that purpose, plants expressing differentially tagged versions of both enzymes, *myc*-FT11 and HA-FT12, were subjected to subcellular fractionation and transmission electron microscopy (TEM) analysis.

3. Test the occurrence of high mannose N-glycoproteins in *Arabidopsis* chloroplasts and identify plastid N-glycoproteins potentially transported directly from the ER to the chloroplast using or sharing components of the lipid precursors pathway.
   - To test this hypothesis, *Arabidopsis* mutants impaired an ER-to-chloroplast direct transport of *lipid precursors*, were used. Comparative studies of chloroplast protein profiles of Columbia 0 wild type, *tgd1* and *tdg4* mutants, disrupted in two key steps of the lipid precursors transport pathway, were performed in order to determine possible proteome differences.
   - Different strategies of chloroplast and stroma fractionation and purification were tested, and further identification of proteins by several methods was also performed.

VII. MATERIAL AND METHODS

1. MATERIAL

1.1. Biological material

1.1.1 *Arabidopsis thaliana* lines:

- Wild type ecotype:*Columbia 0*.
- CAH1(AT3G52720) knock out lineskindly provided by Prof. G. Samuelsson from Umeå Plant Science Centre (UPSC), Sweden:
    - KO:361 (GABI-Kat line 361H01), with a T-DNA insertion in the second intron of CAH1 sequence
    - KO:029(SALK line_029393), disrupted by a T-DNA insertionin the second exon.
- ΔFT or FT11/12 double knock outofcore α(1,3) fucosyltransferases*FT11*(AT3G19280) and*FT12*(AT1G49710) disrupted by T-DNA insertions, corresponding to crossed SALK lines_087481 and 06335 respectively(from Forth et al., unpublished). ΔFT is unable to catalyze the addition of a fucose residuein α(1,3) position to N-glycans.
- *Complex glycan less 1-1 (cgl1)* mutant. Knock out in N-acetylglucosaminyltransferase I(GnTI) (AT4G38240)due to a punctual mutation in G1231A site, leading to accumulation of N-glycoproteins only harbouringoligomannosidic structures but not complex N-glycans (von Schaewen et al., 1993). Kindly provided by Prof.A.von Schaewen, Münster University, Germany.
- FT11-myc/FT12-HA double tagged line. Both genes coding for core α(1,3) fucosyltransferase is fused to an epitope tag on a Columbia 0 background. Kindly provided by Prof. G. Samuelsson.
- Trigalactosyldiacylglycerolmutants: *tgd1-1(tgd1)* (AT1G19800), which contains a nucleotide substitution C/T in an exon, and *tgd4-2(tgd4)* (AT3G06960), a SAIL_760_F05 insertion line at nucleotide 1908. Mutants impaired in lipid precursors import from the ER to the chloroplast.*tgd1-1*is a homozygous line, impaired in in a subunit of the so-called TGD traslocon of the chloroplastinner envelope; *tgd4*, disrupted in an endoplasmic reticulum resident protein has to be maintained asa heterozygous line, due tohomozygous sterility plants.Pool of F_1 offspring seeds of homozygous plants were selected and confirmed by PCR prior to any experimental analysis. Kindly provided by Prof. C. Benning, from Michigan State University, USA.

- Additionally, *Arabidopsis thaliana* ecotype *Landsberg erecta*c ell suspension cultures were used in the experiments performed in UPSC (Burén *et al.*, 2011).

1.1.2 *Nicotiana benthamiana:*

- Wild type *Nicotiana benthamiana* plants grown for 5 weeks were used for leaf transient expression experiments.

1.2. Growth conditions

- *Arabidopsis thaliana* was grown in chambers at 22ºC/18ºC short photoperiod (8 h light/16 h darkness) with a photon flux density of 100 µmol photon/m^2s. Plants were grown on soil for 4-6 weeks, watered with tab water and fertilized with "Combo" Universal Fertilizer (7% total nitrogen, 5% phosphoric anhydride (P_2O_5), 6% potassium oxide (K_2O) and microelements) every two weeks.
- For seed propagation, *Arabidopsis thaliana* (wild type, mutant and stably transformed lines), and *Nicotiana benthamiana* wild type and transiently transformed plants were grown at 24ºC, with long photoperiod day/night cycle (16 h light/8 h darkness).

1.3. Agar plate assays

- *Arabidopsis* root growth assays were performed in Petri dishes containing: 4,40 g/L of Murashige and Skoog (MS) (Duchefa), 1% agarose, 1% sucrose, dH_2O, pH 5,8; autoclaved. Seedlings were grown in theshort photoperiod conditions previously described.
- Before sowing, seedlings were surface sterilized ten times for 1 minute (min) in a 70% ethanol and 0,02% Tween 80 solution, with a final wash with 90% ethanol, followed by 4 washes with freshly autoclaved dH_2O.
- Different volumes of a sterile 1M NaClstock were added to achieve the different final concentrations (0, 50, 75, 100 and 150 mM) in the salt stress analysis.
- Pictures of the plates were taken with a Nikon digital camera with polarized light, ISO 100 and CWB 3000.

2. METHODS

2.1. Subcellular fractionation techniques

2.1.1. Total chloroplast and stroma isolation

The method is based on Kunst's protocol (1998) with modifications by Seigneurin-Berny et al. (2008). Briefly, 20 g of *Arabidopsis* leaves were placedin ice cold water for 20 min prior to homogenization, which was performed by two burst of 2 s each, at full polytron speed (Waring) in 300 mL of cold homogenization buffer. The homogenate was filtered trough 2 gauze layers and nylon mesh (80 µm). The filtrate was centrifuged for 90 s at 1465*g* at 4ºC. Crude chloroplasts pellet was carefully resuspended in ice cold resuspension buffer (RB) and loaded on a cold Percoll gradient (performed by mixing 30 mL of cold 100% Percoll and 30 mL of cold 2xRB, and centrifuging 1 h at 26.900 *g*), and centrifuged for 10 min at 1465*g*. The band corresponding to intact chloroplasts was collected and diluted 3 times with RB, and centrifuged 1 min at 1.465 *g*. The pellet containing the total chloroplasts fractionwas resuspended in 20 mM4-(2-hydroxyethyl)-1-piperazineethanesulfonic acid (HEPES) buffer, pH 7.5 containing 10 mM $MgCl_2$ which leads to an osmotic shock and chloroplasts burst. After that, samples are mixed by vortexing and sonicated. To obtain stroma fractions, total chloroplast suspensions were centrifuged 20 min at 14.000*g*. The supernatant contains soluble stroma proteins. Protein or chlorophyll concentration was determined as described below and samples were stored at -20ºC until use.

2.1.2. Microsome isolation

Arabidopsis leaves were grinded with liquid nitrogen and then 15 mL of homogenization buffer (25mM Tris-HCl pH 7,5, 0,25 M sucrose, 3 mM EDTA, 1 mM DTT) and 1x PIC was added. Samples were vortexed and centrifuged at 10.000 *g* in a fixed angle rotor for 15 min at 4ºC. Supernatant was filtered through a mesh of 15 µm diameter pore. Filtered material was centrifuged at 150.000 *g* 1h 10 min. The pellets (total microsomes) were resuspended in 500 µL of Buffer A(5 mM Tris- HCl pH 7,5, 0,25 mM sucrose, 3 mM EDTA, 1 mM DTT) and kept for further analysis.

2.1.3. Microsome sucrose gradient

Sucrose gradient separatesmicrosome vesicles of different densitiy, which is useful for estimating the localization of a protein of interest within the endomembrane system.Sucrose gradients were composed by equal amounts of a dense phase (50% (w/v) sucrose in buffer B (5 mM Tris-HCl pH

7,5, 3 mM EDTA, 1 mM DTT)) and a lighter phase (15% (w/v) sucrose in buffer B). Both phases were carefully placed in an assay tube, and incubated 3h at 4ºC in horizontal position. Microsome fractions were frozen, thawed and homogenized (process that favour separation of the material) and 500 µL were loaded onto the sucrose gradients. Gradients were centrifuged from 5 h to overnight in swing-out rotor at 80.000 *g* at 4ºC. After that, 500 µL aliquots of the gradients were recovered from the top to the bottom of the tube and subjected to protein precipitation.

2.1.4. Protoplast isolation from *Arabidopsis thaliana* mesophyll cells

Modified from Sheen *et al.* (1997). Briefly, about 10 plants of a desired line were selected, and leaves were cut into 1 mm strips. Strips were then digested for 2-3 hours with 40 rpm agitation, in 60 mL of enzyme solution (0.5 g celulase Onozuka R5 (Duchefa) (1% (p/v)), 0.1g macerozyme R10 (0.2 % (p/v)(Duchefa)), 25 mL 0.8 M Mannitol (0.4 M), 2 mL 0.5 M KCl (20 mM), 2 mL 0.5 M MES pH 5.7 (20 mM)), heated to 55ºC 10 min, addition of 0.5 mL $CaCl_2$ (10 mM), 0.05 g BSA (0.1 % (p/v)), 20.5 mL distilled water and subsequently filtered through a 0.45 µm filter). The solution containing the released protoplasts was filtered trough a 41 µm-pore diameter nylon mesh. Filtrated material was centrifuged 2 min at 100 *g*. Pellet containing intact protoplast was washed with 5 mL of W5 solution (154 mMNaCl, 125mM $CaCl_2$, 5 mMKCl, 2 mM MES pH 5.7). Samples were then centrifuged again 2 min at 100 *g.* Thepellet was resuspended in 5 mL of MMg solution (0.4 M Mannitol, 15 mM $MgCl_2$, 4 mM 0.5M MES pH 5.7). Protoplasts were then counted using a Neubauer chamber in order to estimate the amount of cells per mL to calculate dilutions for transfection or culture optimal concentration (about 4.10^5 cells/mL). In protoplast liberation assay realeased protoplasts were counted at 2, 4 and 6 hours after enzyme solution addition.

Liquid media for protoplasts culture: W5, W5a (5 mM KCl, 2 mM MES pH 5,7, 125 mM $CaCl_2$, 154 mM NaCl, 10 mM glycose and 1 µM NAA) and Medium B (0,4 M Mannitol, 10 mM glucose, 1µM NAA, 3,2 $g.L^{-1}$ Gamborg Media, pH 5,8)(modified from Schirawski *et al.*, 2000)

Protoplast isolation from *Arabidopsis* cell cultures were performed in a collaborative work with Prof. G. Samuelsson's group in UPSC (Sweden), as described in Burén *et al.*, 2011.

2.2. Molecular biology techniques

2.2.1. Genomic DNA isolation from *Arabidopsis thaliana* leaves

About 100 mg of leaf tissue was was grinded in 400 µL of extraction buffer (200 mM Tris-HCl pH 7,5, 250 mMNaCl, 25 mM EDTA, 0,5% SDS, dH_2O), and incubated 10 min at 65ºC. Then 130 µL of KAc 3 M pH 5, was added, mixed and incubated 5 min at -20ºC. The sample was

centrifuged 10 min at 14.000 *g*, and 450 µL of SN were transferred to a new tube with equal volumeof isopropanol, inverted and incubated 5 in at room temperature. Then it was centrifuged 10 min at full speed, and the supernatant removed. 500 µL of 70% EtOH were added to the remaining pellet, and then centrifuged 5 min at full speed. EtOH was removed and pellet was dried and resuspended in 50 µL of TE Buffer (10 mM Tris-HCl pH 8, 1 mM EDTA).

2.2.2. Plasmid DNA isolation

Plasmid DNA was isolated from 3-5 mL *Escherichia.coli* or *Agrobacterium tumefaciens* cultures, grown from single colonies in LB media (5 g/L yeast extract, 10 g/L Triptone, 5 g/L NaCl) with the proper antibiotic, over night at 37ºC for *E.coli*and 48h at 28ºC and darkness for *Agrobacterium*. Extraction was performed following manufacturer's indications (EZNA, Promega).

2.2.3. RNA isolation

Arabidopsis seedlings were grinded to fine powder, and homogenized in 1 mL of TRIzol reagent (Invitrogen) per 50-100 mg of tissue.Then samples were incubated 5 min at 4ºCand centrifuged at 15.000 *g* at 4ºC. Supernatants were mixed with same volume of chloroform. Samples were mixed, incubated 2-3 min at 4ºC, and centrifuged 15 min at 13.000 *g*. RNA was precipitated by mixing the obtained aqueous phase with 500 µL of isopropanol per 1 mL TRIzol reagent, and incubating for 10 min at room temperature afterwards. Then samples were centrifuged 15 min at 13.000 *g* at 4ºC. For washing the RNA, pellets were washed with 1 ml of 75% ethanol per 1 mL TRIzol reagent, samples were centrifuged for 5 min at 13.000 *g* at 4ºC. Pellets were air-dried 10 min at room temperature, and resuspended in 40 µL of DEPC water (Merk). Samples were then incubated 10 min at 65ºC and left from 30 min at 4ºC and then centrifuged for 10 min at 13.000 *g* at 4ºC. The supernatant contains the isolated RNA, ready for further analysis.

2.2.4. DNA and RNA measurement

DNA and RNA samples were measured using NanoDrop (ND 1000, Thermo Scientific™) device, following manufacturer's instructions. 1 µL of sample was loaded into the detector, blank was dH_2O, and absorbance was measured at 260 nm and 280 nm, using the NanoDrop software.

2.2.5. Coding DNA (cDNA) synthesis

cDNA synthesis was performed using kit iScript™ from BioRad, following manufacturer's instructions.Briefly, 300- 500 ng of RNA template (isolated as described in section 2.2.3) were mixed with 4 µL of 5x iScript reaction mix, 1 µL iScript reverse transcriptase and DEPC water to a final volume of 20 µL. Reverse transcription reaction was performed as follows: 5 min at 20ºC, 30 min at 42ºC, 5 min at 85ºC and hold at 4ºC. 2 µL of the cDNA were used for PCR amplification.

2.2.6. Polymerase Chain Reaction (PCR)

DNA was amplified by PCR. Primers used are described in the following sections. All standard reactions were performed with isolated DNA and using Biotools reagents in the following proportions: 0,5 µL dNTPs (10mM), 2,5 µL 10x Reaction Buffer, 0,75-1 µL $MgCl_2$ (50 mM),0,75-1 µL Primer Fordward (10 µM), 0,75- 1 µL Primer Reverse (10 µM), 0.75 -1 µL DNA Polymerase (1U/ µL), dH_2O and DNA sample toa final volume of 25 µL.Standard PCR program used was: Denaturing step, 94ºC 5 min and 30 s; Annealing step, specific Tm 30 s; Elongation step, 72ºC 30 s-1 min, and 5-7 min, Maintenance, 14ºC undefined time. Denaturing, annealing and elongation steps were repeated 30-35 cycles. Also temperature gradient reactions were performed prior to test every primer pair.For specific cloning reactions when high accuracy was requiredhigh fidelity polymeraseswith lower error rate during replicationwere used: *Pfu* kit (Fermentas and Biotools) and *Pfx* kit (Invitrogen), following manufacturer's indications. For all reactions Termocylcer MyCycher™ (BioRad) was used.

2.2.7. Generation of point mutated versions of CAH1 (directed mutagenesis)

- Cloning of HA-tagged CAH1mutant variants

Mutated glycoforms were generated from a HA-tagged version of CAH1 (HC), which was included in a pPE1000 NanoT plasmid (Hancock, 1997). Site-directed mutagenesis using a Phusion Site-Directed Mutagenesis Kit (FINNZYMES OY) was performed. Single mutations (N1, N2, N3, N4 and N5) were created performing PCR by using the 5'-phosphorylated primer pairs N1for and N1rev, N2for and N2rev, N3for and N3rev, N4for and N4rev, N5for and N5rev, respectively and the tagged HA version of CAH1 (HC) included in the plasmid (pHACAH1) as a template. Double mutations (N1+N2, N3+N5, N3+N4 and N4+N5) were created using pHACAH1-N1 with primer pair N2for and N2rev, pHACAH1-N3 with primer pair N5for and N5rev, and pHACAH1-N4 with primer pairs N3for and N3rev or N5for and N5rev, respectively.

Triple mutant N3+N4+N5 was created using pHACAH1-N4+N5 with primers N3for and N3rev. Quadruple “Q” mutants and the non-glycosylated HA-tagged CAH1 “NG” were created by cutting single, double and triple mutants with NcoI and SacI, respectively, and ligating vector backbone and inserts to create Q1 (N2 backbone and N3+N4+N5 insert), Q2 (N1 backbone and N3+N4+N5 insert), Q3 (N1+N2 backbone and N4+N5 insert), Q4 (N1+N2 backbone and N3+N5 insert), Q5 (N1+N2 backbone and N3+N4 insert) and NG (N1+N2 backbone and N3+N4+N5 insert).Cysteine mutant clones were created using PCR with primers that introduce the restriction site corresponding to XhoI. Single mutants C1, C2, C3 and C4 were generated by ligating restricted PCR fragments amplified using primers C1for and C1rev, C2for and C2rev, C3for and C3rev, and C4for and C4rev, respectively. The C1+C3 double mutant was created by cutting the C1 and C3 mutants with NcoI and BsrGI and ligating the C1 vector backbone with the C3 insert fragment.C terminus mutant clones were created using PCR primers that introduce the restriction site corresponding to XhoI. In the case of C terminus deletion (Cdel), XhoI restriction site was placed 48 nucleotides upstream than C terminus “Cmut” mutant. Then, “C mut” mutant was generated by ligating restricted PCR fragments amplified using primers Cmut for and Cmut rev, while “Cdel” mutant was generated following the same method, but using primers “Cmut for” and “Cdel rev” instead.DNA sequences of all constructions were verified by DNA sequence analysis and/or restriction analysis to ensure they had the correct sequences and reading frames.

- Primers used for cloning of HC variants (5’-3’)

  - ForCAH1/KpnI TTACAGGTACCATGAAGATTATGATGATGA
  - RevCAH1/SacI ACTTTGAGCTCAAATGTTTGAACGAGAATT
  - N1for TACTACTTCACAGCCGCAACACTAGTGAA
  - N1rev TTCACGGTGTATTGAATTCAATTTG
  - N2for ATAGAAAACAAGGCCTATACCTTACTGCAAAT
  - N2rev TATCACATCTCCTGCTCCCTCCC
  - N3for AGACTCAAAGGGGCCCACACAGCACAAG
  - N3rev CTCTTCCTTTAGCTTCACCAATTTCTC
  - N4for CCTTGCTCCGAGGCCGTTTCTTGGAC
  - N4rev AGGAGTAGTGAGTGAACCAATGTATCT
  - N5for ACTTCTTTCAAGGCCAATTCAAGACCGT
  - N5rev GTCCAATGGAGATCTGAGTAGTTCTAC
  - C1 for ACACTCGAGGTCGGTAAATTGCAATCTCCA
  - C1 rev GACCTCGAGTGTGGTGAAGTGAGGGTTTAA
  - C2 for CGTCTCGAGTGTTGCCATGTTCTTCGGGGA
  - C2 rev ACACTCGAGACGTGGTTCACTAGTGTTGCG

- C3 for TCCCTCGAGCGAGAACGTTTCTTGGACCAT
- C3 rev TCGCTCGAGGGAGGAGTAGTGAGTGAACCA
- C4 for ACCCTCGAGACCCCTCAACGGCCGGAGAGT
- C4 rev GGTCTCGAGGGTCTTGAATTGTTCTTGAAA
- Cmut for ATACTCGAGAAACCCAATTAAGGATCCACT
- Cmut rev TATCTCGAGTTTGTTACCGGTTTCTTTTTT
- Cdel rev GTGCTCGAGCYCAACTCTCCGGCCGTTGAG

- Primers used for cloning the variants into a binary vector (pmDC32) (5'-3')

- ForCAH1/KpnI TTACAGGTACCATGAAGATTATGATGATGA
- RevCAH1/SacI AATTCTCGTTCAAACATTTGAGCTCAAAGT
- RevCAH1**b**/SacI GCAGATCGTTCAAACATTTGAGCTCAAAGT

- Primers used for sequencing

- ForCAH1/KpnI TTACAGGTACCATGAAGATTATGATGATGA
- For 35S-For GCAAGACCCTTCCTCTATA
- RevTnos-begin CCTGTTGCCGGTCTTGCGAT
- RevTnos AAACTTTATTGCCAAATGTTTGAACGA

2.2.8. PCR screening for TGD4-2 allele

It order to identify TGD4-2 allele, with a T-DNA insertion, two standard PCR reactions, one amplifying *tgd4* gene using LP and RP specific primer pair, and other amplifying T-DNA insertion using LB and RP primer pair were performed in each sample.

- Primers used for screening (5'-3')

- **LP:**TCCCTTCTATTGAAACATTCTTCTG
- **RP:**TGAGTTCTCCAATTCCAATATTTG
- **LB:**GCCTTTTCAGAAATGGATAAATAGCCTTGCTTCC

2.2.9. Reverse Transcriptase Polymerase Chain Reaction (RT-PCR)

RT-PCR is a very sensitive technique used for measuring gene expression (Powell *et al.*, 1987; Wilkinson *et al.*, 1998). Reaction is performed in two steps. First, coding DNA (cDNA) is synthesized from RNA template as previously described (see section 2.4.4). Secondly, standard PCR is performed, with 2 μL of cDNA and performing 1 min for time extension. Control reactions were performed without reverse transcriptase.

- Primers used for RT-PCR (5'-3')
  - Endo β (1,4) glucanase (*kor1*, AT5G49720)
    - KOR1 Fordward: TGCCTTCTGGGGTGGCCCTT
    - KOR1 Reverse: GCAGCAGCGACAGAATCACACAA
  - ADP rybosylation factor (*adp-rf*, constitutive gen. AT3G62290)
    - ADP-RF Fordward: GGTCTGATGCAGCTGGTAAGACTAC
    - ADP-RF Reverse: TGTTAGAGAGCCAGTCAAGTCCCTCA

2.2.10. Agarose gel electrophoresis

DNA fragments were separated by electrophoresis in horizontal 1% to 2% w/v agarose gels. Agarose concentration varied according to the size of the fragment subject of experiment. TAE buffer (40 mM Tris-HCl pH 7.6, 20 mM acetic acid and 1mM EDTA) (Sambrook, 1989), was used both for diluting agarose and performing the electrophoresis. Samples were mixed with loading buffer (80% glycerol, 0.5M EDTA pH 8.0, 35% dH_2O and 25 mg of bromophenol blue). Size of obtained fragment was measured by comparing its relative electrophoretic mobility to standard markers of molecular weight (1 Kb and 100 bp ladder, Biotools).

2.2.11. Isolation and purification of DNA fragments

DNA fragments of the appropriatesize were carefully excised from agarose gels and purifiedusing Promega or E.Z.N.A kits following manufacturer's inditations. When possible PCR product was directly purified for further analysis (i.e., restriction digestion, see section 2.2.12) by using a specific kit for liquid PCR samples (Qiagen, Promega) following manufacturer's guidelines.

2.2.12. Restriction digestion

Genomic and plasmid DNA digestion with restriction enzymes were performed according to manufacturer's instructions (Takara, Fermentas). For cloning of HA-tagged CAH1 mutant variants into the pMDC32vector, the DNA sequences contained in pPE1000 vector and amplified with specific primers For CAH1/KpnI and Rev CAH1/SacI, using the high fidelity *Pfu* polymerase were digested with restriction enzymes KpnI and SacI (Takara). pMDC32 vector was also digested with the same enzymes, in order to ligate the end in the following step. Plasmid DNA digestions were also performed for further checking the fragment of interest was correctly inserted in the vector.

2.2.13. Ligation and cloning reactions

Restriction digested DNA fragments ligated tothe proper vector (pPE1000 or binary vector pmDC32) using T4 ligase enzyme (Biotools, Promega or Fermentas), following manufacturer's instructions. For cloning into pMDC32 vector, the reaction mixture was incubated at 18ºC overnight. The ligate products were then sequenced, and subsequently introduced into E.coli.

2.2.14. Generation of competent *E.coli* DH5α cells

E.coli DH5α Competent cells were prepared according to $CaCl_2$ classic method (Kushner, 1978) with modifications: culture wasgrown either from glycerol preparations or Petri dish plates. A 4,5 mL pre-inoculum was put in LB with ampicillin (our bacteria has resistance against this antibiotic) andgrown overnight at 37ºC in shaker. 2,5 mL were added to 200 mL fresh medua, and culture was growth at 37ºC in continuous shaking until it reached DO_{600}=0,4-0,5. Then the culture was centrifuged at 1.700 *g* 10 min at 4ºC, and the pellet was resuspended in 30 mL of TFB-I (30 mM AcK pH 5,8, 50 mM$MnCl_2$, 10 mM$CaCl_2$, 100 mM RbCl, 15% glycerol) and incubated at 4ºC for 10 min. Afterwards, culture was centrifuged at 1.700 *g* for 10 min, and pellet was resuspended in 4 mL of TFB-II (10 mM MOPS pH 7, 75 mM $CaCl_2$, 10 mM RbCl, 15% glycerol). Then 200 µL aliquots were stored at -80ºC.

2.2.15. Heat shock transformation of competent *E.coli* DH5α cells

10 µL of ligation product (containing vector with desired DNA fragment inserted) were added to 200 µL of *E.coli*DH5α cells. Mixture was incubated 30 min in ice, and heat shock was produced 2 min at 42ºC, followed by 1 min in ice again, increasing cell membrane permeability for DNA incorporation. After that, 800 µL of LB were added and culture was incubated 1h at 37ºC under

250 rpm agitation. Different concentrations of the grown cultures were dispensed in solid LB media with proper antibiotics, and incubated at 37ºC overnight. Transformed single colonies were checked for the presence of the insert by colony PCR using specific primers for each insert.Then plasmid DNA of positive colonies was purified using a Promega plasmid DNA isolation kit. All plasmids were then sequenced and stored in 80% glycerol at -80ºC.

2.2.16. Electroporation of electrocompetent *Agrobacterium tumefaciens* cells

15 µL of plasmid DNA sample, with a concentration of 1-5 µg/µL were added to 40 µL competent *Agrobacterium* cells (GV3101 containing pM90), on ice. Mixture was transferred into a cold electroporation cubette (BioRad) and subjected to 2,5 kV electric current in a GenePulser (BioRad) device. Then 1 mL LB was added, and transferred to an eppendorf tube incubated for 3,5h at 28ºC in a shaker. Afterwards, samples were centrifuged at 13.000 *g* and the pellet was spread on LBwtih proper antibiotic resistance (rifampicyn, gentamicyn and kanamicyn)and incubated at 28ºC for 48h. Presence of the insert was checked by colony PCR.

2.3. Protoplast and plant tranformation

2.3.1. Protoplast transfection with polyethilenglycol (PEG)

Protoplast from wild type *Arabidopsis* leaves were transfected with tagged CAH1 versions cloned in pPE1000 plasmid.10 µl of plasmid DNA at 1-2 µg/ µl concentration were added to 100 µL media containing around 4.10^6 cells/µL. Afterwards, 110 µL of PEG 4000 was added to ease the entry of DNA into the cells. Mixture was incubated at room temperature during 15 min. Then 440 µL of W5 solution (described in section 2.1.4) were added and samples were centrifuged at 100 *g* 3 min, PEG was removed and protoplast were resuspended in 500 µL of W5 solution. Samples were incubated in multi well plates at room temperature in darkness for 24h. Trasfected protoplasts were then recovered and centrifuged at 90 *g* 5 min. The pellet (isolated protoplasts) were resuspended in 50 µl of KEB buffer and stored at -20ºC for further analysis.

2.3.2. *Nicotiana benthamiana* transformation by agroinfiltration

Modified from Sparkes et al. (2006). Selected colonies of *Agrobacterium* transformed with the desired inserts in pMDC32, and colonies containing the p19 plasmid to avoid gene silencing (Voinnet et al., 2003) were grown in 10 mL LB or YEB (0,5% w/v beef extract, 0,1 w/v yeast extract, 0,5 w/v peptone, 0,5 w/v sucrose, 1,5 w/v agar, 0,048 w/v $MgSO_4$) and cultured, at 28ºC in darkness and shaking for 24-48h. Grown cultures were refreshed in order to reach exponential

growth phase of the bacteria. They were then centrifuged 5 min at 4500 *g* and pellet was resuspended in 2 mL of agroinfiltration buffer (10 mM MES, 10 mM $MgSO_4$, 0,2 mM acetosyringone, pH 5,7), incubating 3h at room temperature in darkness. OD_{600} was then measured in Biophotometer (Eppendorf) and dilutions with agroinfitration buffer were performed in order to have a final of 0,5. Subsequently, the agroinfiltration solution was prepared, by mixing equal volumes of 0,5OD_{600} p19 culture + pMDC32 with desired inserts. Infiltration is carried out in 1mL syringe, in the underside of selected leaves (fully developed). 2-3 leaves per plantwere infiltrated and transformed plants were covered with transparent film in darkness overnight. Finally, plants were moved to culture chamber for two days, before sample extraction.

2.3.3. *Arabidopsis thaliana* flower dipping

Arabidopsis KO:29 and KO:361were stably transformed with*Agrobacterium* cultures. Protocol was adapted from Clough and Bent, (1998) with modifications of Bechtold et al.(1993). KOs *Arabidopsis* mutants were grown in long light cycle chambers until flowering. Meanwhile, *Agrobacterium* cultures (carrying pMDC32 –Cdel, Cmut, NG-) were growth as described in 2.3.2 section, to a final volume of 30 mL. Culture wasthen centrifuged, and OD_{600} adjusted to 0,8, by diluting in 5% sucrose solution and0.05% of Tween 20 before dipping. Inflorescences the closed flowers were dipped in the *Agrobacterium* solution. Plants were maintained in high humidity from 24h. Dipping was repeated afterone week in order to increase transformation rate. Harvested seeds are subjected to further selection.

2.4. Protein and electrophoresis techniques

2.4.1. Protein extraction (TE)

For total protein extract (TE), *Arabidopsis* leaves were grinded with liquid nitrogen. The powder was resuspended in kinase extraction buffer, KEB(25mM Tris-HCl, 10mM $MgCl_2$, 5mM ethylene glycol tetraacetic acid (EGTA), 2mM dithiotritol (DTT), 10% (v/v) glycerol, 75 mMNaCl, 0,2% (v/v) Igepal CA630 and 1x Protease Inhibitor Cocktail (PIC) (Complete, Roche)). After vortexing the mixture, samples were centrifuged 20 min at 4ºC at 13,000*g*. Supernatant was kept at -20ºC for further analysis.

Total extract from *Nicotiana* leaves was isolated 2 days after agroinfiltration. Leaf tissue was grinded in a cold pestle kept on ice, with 500 µL KEB buffer, centrifuged at 13000 *g* at 4ºC for 30 min and the supernatant was kept for further analysis.

2.4.2. Protein measurement

- *Bradford:* Protein concentration of leaf extracts was measured following the method described by Bradford (1976) with two different kits. In the first, Bradford protein reagent (Biorad), 800 µL of sample was mixed with 200 µL of reagent, samples were shaked and incubated 15 minutes in darkness. Afterwards, O.D was measured at 595 nm. The standard curve was prepared was determined by measuring known concentrations (0 to 2 mg/mL) of bovine serum albumin (BSA) (Merk). The second, the *BCA Protein assay* (Thermo Scientific^{TM}) was used in Münster´s lab, following manufacturer's instructions (briefly, BCA solution was prepared and added to both standard and samples in a plate, mixed and incubated at 55ºC for 30 min). Absorbance at 562 nm was measured.
- *RC-DC^{TM} Protein assay*: Protein measurement for DIGE was performed using RC-DC^{TM} Protein Assay (BioRad) following the manufacturer's instructions. Briefly, standard curve with known concentrations of BSA from 0,2 to 1,5 mg/mL was diluted in lysis buffer (see section 2.4.5). 1/10 dilutions of samples were prepared in lysis buffer: 125 µL of Reactive I was added to 25 µLof each standards and samples, and reaction was incubated for 1 min. Then, 125 µL of Reactive II was added and samples were mixed and centrifuged at 14000*g*for 5 min. Supernatant was discarded and 127 µL of Reactive A' were added to the dried pellet,mixed and incubated for 5 min. Finally 1 mL of Reactive B vas added and mixed. Sample was incubated 15 min and absorbance at 750 nm was measured.

2.4.3. Protein precipitation

- *Acetone method:*1 mL of pure cold acetone was added to 200 µL sample, and incubated from 2h to overnight at -20ºC. Samples were then centrifuged at 4ºC at 15000*g*, and pellet was dried, ready for further analysis.
- *DOC-TCA*: DOC (Na deoxycholate, detergent) (0,15% v/v) was added to the protein sample, vortexed and incubated 10 min at RT. Then, 72% TCA (trichloroacetic acid) (0,13% v/v), was addedand vortexed, and samples were incubated 3h in ice, centrifuged 20 min at15000 *g*at 4ºC. Pellets werecarefully washed with 300 µL of acetone, incubated again 20 min in ice, and centrifuged another 20 min at 15000 *g* at 4ºC. Dried pellets were resuspended in protein loading buffer (see 2.4.2).

2.4.4. 2-D Fluorescence Differential Gel Electrophoresis (2-D DIGE)

- Sample preparation and protein labeling

Stroma extracts were first purified in order to separate proteins from detergents, salts, lipids, phenolics, nucleic acids, and other interfering contaminants, by using 2-D Clean Up kit (GE Healthcare) following manufacturer's instructions. Subsequently, obtained pellets were resuspended in lysis buffer (30 mM Tris-HCl pH 8,5, 7 M Urea, 2 M Thiourea, 2% CHAPS) to a final protein concentration of 5 μg/μL.Samples were then labeled with N-hydroxysuccinimidyl ester-derivates of the cyanine dyes Cy2, Cy3 and Cy5 (Amersham Biosciencies) at 4ºC and darkness. 50 μg of sample were minimally labeled with 400 pmolof either Cy3 or Cy5 and a pool of bothsamples was labeled with Cy2, as an internal standard for image and cross-gel statistical analysis.Then, reaction was stopped by adding excess of L-lysine 10 mM (Sigma) for 10 min on ice. Labeled samples were mixed and the final volume was adjusted to 150 μL with lysis buffer; volume of the Cy3 and Cy5 labeling reactions (corresponding to 50 μg each) were mixed with an equal amount of Cy2 labeled standardand run on the same gel. Four replicates, alternating the samples stained with the different dyes (dye swap) in order to avoid possible artifacts were studied.

- Protein separation by 2D Gel Electrophoresis and Gel Imaging

Immobilized non-linear pH gradient (IPG) strips, pH 4-7 (Amersham Biosciences, Inc), were rehydrated in rehydration buffer (8M Urea, 2% w/v CHAPS, 0,002% of bromophenol blue), containing 18 mM DTT and 12,5 μL of IPG buffer specific for pH 4-7, in darkness at room temperature overnight, according to manufacturer's guidelines. Briefly, strip was placed in arehydration tray (Amersham Biosciences, Inc), submerged in the buffer (acrylamide surface upside down) and covered with cover fluid (GE Healthcare) avoiding bubble formation. Rehydration buffer containing 180 mM DTT was added to each sample and samples were pooled and incubated 10 min before loading onto the rehydrated strip.Isoelectric focusing was performed using an Ettan IPGphor3 (GE Healthcare)for a maximum of10 kV-h,50 μA during 7-8h. Strips were then equilibrated for 15 min in equilibration buffer(50 mM Tris-HCl, pH 8,8, 6 M Urea, 30% (v/v) glycerol, 1% (w/v) SDS) adding 100 mM DTT), and then for 15 min in the same buffer adding 250 mMiodoacetamide (IAA). Equilibrated IPG strips were transferred onto 10% acrylamide gels placed between low fluorescence glass plates (GE Healthcare). Strips were covered with 0,5% (w/v) a solution low melting point agaroseprepared in running buffer containing bromophenol blue. Gels were run in Protean II gel tanks (Bio-Rad) at 30 mA per gel at room temeperature. Ready 2D gels were immediately sent to Proteomics Service Facility in

Universidad Complutense - Parque Científico de Madrid (UCM-PCM) (Madrid), where they werescanned using a fluorescence scanner Typhoom 9400 (GE Healthcare), the excitation/emission filters corresponding to each fluorocrome (Cy3: 532/580, Cy5: 633/670 y Cy2:488/520 nm wavelength). High resolution images (pixel size: 100 µm) were acquired for further protein profile analysis.

- Image analysis, Gel staining, and Spot Picking

This analysis was carried out in the Proteomics Facility Service UCM-PCM (Madrid). Briefly, images were processed using software ImageQuant v5.1 (GE Healthcare) to eliminate areas lacking protein signal and then further analyzedusing DeCyder v6.5 software (GE Healthcare). For DeCyder image analysis, the differential in-gel analysis (DIA) mode of DeCyder was first used to merge the Cy2, Cy3, and Cy5 images for each gel and to detect spot boundaries for the calculation of normalized spot volumes/protein abundance. Spot co-detection was performed with the proper software algorithm, specifying the detection of maximum spots as possible. Spots corresponding to backgroundwere filtered out (by software and manually if necessary). Then, in was used the biological variation analysis (BVA) mode of DeCyder, which takes into account all biological replicates of the experiment.It matches all pairwise image comparisons from difference in-gel analysis for a comparative statistical analysis. Student's t test p-value was calculated using the values from each experimental condition. Gels were stained following Comassie colloidal protocol, in order to visualize total protein content.Spots of interest were excised from 2D gels using Commassie colloidal stained image as reference, by Ettan Spot picker robot (GE Healthcare) following the manufacturer's instructions.Spots were transferred to microwell plates and kept frozen at −20°C for protein identification by MALDI-MS.

2.4.5. Denaturing polyacrylamide gel electrophoresis (SDS-PAGE)

Proteins from leaf total extract, protoplasts, chloroplasts and stroma fractions were separated by SDS-PAGE, to determine their characteristic protein pattern.Samples were prepared by mixing with an equal volume of 2x protein loading buffer (80% glycerol, 10% SDS, 0,5 M Tris pH 6,8, 5% β mercaptoethanol, 0,0025% bromophenol blue), heated 5-10 min at 95ºC and centrifuged.Electrophoresis was performed in vertical acrylamide/bisacrylamide gels in denaturing conditions (Laemmli, 1970). Acrylamide concentration of the separating gel varied from 10% - 12%, according to the separation of bands required within the pattern.Electrophoresis was performed in Mini Protean III (BioRad) device following manufacturer's indications. The molecular mass of protein bands was estimated using standard protein marker (Prestained SDS-PAGE Standard Broad Range or Dual Color) (BioRad) as a reference. Gels were then stained with

Coomasie Blue or silver (Silver Staining kit PlusOne, Amersham, GE), or alternatively transferred to nitrocellulose membranes for immunodetection.

2.4.6. Immunodetection assays (Western Blot)

Acrylamide gels were washed 15 min in transfer buffer (25 mM Tris, 192 mM glycine, 20% methanol), in order to remove SDS excess. Proteins in acrylamide gels were electrotransferred at 100 V during 1 h, onto a nitrocellulose membrane (Membrane Blotting, Pall Corporation) for further immunological detection (Western Blot) or Concanavalin A detection (Affinoblot) (see section 2.4.7). For immnunodetection assay, membranes were blocked 1 h at RT in 5% milk in washing buffer ((0.1 M Tris-HCl pH 7.5, 0.5 M NaCl, 1 % (v/v) Tween-20). After two washing steps of 10 min to remove excess of 5% milk, membranes were incubated overnight in agitation at 4ºC in primary antibody (1:1.000 dilution, see below) in 2% milk washing buffer). Membranes were then washed twice 10 min, and incubated 1 h at RT in agitation with secondary antibody linked to horseradish peroxidase (1:10.000 dilution, see below) in 2% milk washing buffer. After two more 5 min washes one distilled water wash was performed. Immunodetection was performed using a chemiluminescent method, the enhanced chemiluminescence (ECL) system (Amersham) andImmun-Start^{TM}Western C^{TM} Chemiluminescent (BioRad), following the manufacturer's protocols. Images were taken using BioRad Molecular Imager Chemic Doc^{TM} XRS with Image Lab^{TM} software.

Primary antibodies used:

- Polyclonal rabbit antibody against αCAH1 protein from *Arabidopsis thaliana.* It was used in order to check the presence of the protein in the different isolated fractions, and as control in order to check the proper functionality on endomembrane route.Kindly provided by Prof. G. Samuelsson in UPSC (Sweden).
- Polyclonal rabbit antibody against BiP protein from *Arabidopsis*. Is a chaperone resident in ER (Pedrazzini et al, 1997) and was used as control of ER contamination in fractionation experiments, and for measuring co-precipitation with tagged CAH1 versions. Kindly provided by Prof. G. Samuelsson in UPSC, (Umea, Sweden).
- Polyclonal rabbit antibody against core α(1,3)fucoseresidues. It was used for measuring fucosylation levels in *Arabidopsis* samples.Kindly provided by Prof. Patrice Lerouge, form Rouen University (France).
- Polyclonal rabbit antibody against core β(1,2)xylose residues. It was use for measuring xylosylation levels in *Arabidopsis*samples. Kindly provided by Prof. Patrice Lerouge, form Rouen University (France).

- Polyclonal rabbit antibody against Ribulose-1,5-bisphosphate carboxylase oxygenase (RuBisCo). It was used as chloroplast loading control. Kindly provided by Prof. G. Samuelsson in UPSC (Sweden).
- Monoclonal mouse antibody against hemmaglutinine (HA)tag (Covance). It was used for checking the presence of the HA-tagged CAH1 constructs, HA- tagged FTs constructs (both in Western Blot assays and immunocitochemistry for TEM) and to measure the amount of protein that co-precipitate with BiP.
- Polyclonal rabbit antibody against myc tag (Santa Cruz). It was used for checking the presence of myc-tagged FTs constructs, both for Western Blot and immunocitochemistry for TEM.
- Polyclonal mouse antibody against thioglucoside glucohydrolase or myrosynase TGG1, kindly provided by J. Meijer from Uppsala University (Sweden).
- Polyclonal rabbit antibody against vacuolar protease carboxypeptidase Y (CPY).Kindly provided by Prof. E. Rojo from CNB (Spain)
- Polyclonal rabbit antibody against 2-light-harvesting complex 2 (Lhc2) (Agrisera)

Secondary antibodies used in Western Blot experiments were ECL Donkey anti-rabbit IgG,rabbit polyclonal antibody linked to horseradish peroxidase (HRP) (GE Healthcare UK), or mouse (Goat anti-mouse HRP conjugate, BioRad).

2.4.7. Concanavalin A lectin detection (affinoblot)

This technique was used for the detection of high-mannose type N-glycans of glycoproteins, using Concanavalin A (Con A) (Sigma), a plant with high affinity for this kind of oligosaccharide side chains (Faye and Chrispeels, 1985). Nitrocellulose membranes were incubated overnight at 4ºC in Tween Tris-Buffered Saline (TTBS) (10 mM Tris-HCl pH 7.5, 0.5 MNaCl, 0.1% (v/v) Tween 20). Membranes were incubated with 15 mL of Con A Buffer (10 mM Tris-HCl pH 7.5, 0.5 MNaCl, 0.1% (v/v) Tween 20, 1 mM $CaCl_2$, 1 mM $MnCl_2$) and 15 µL of Con A (25 mg/mL stock) for2 h at room temperature. Then 4 washes with TTBS every 15 min during 1 h were done. Membrane was then incubated 1 hour at RT with 15 mL TTBS and 3µL peroxidase (POX) (Stock: 50 mg/mL). POX is a glycoprotein which binds to Con A, and allows detection through chemiluminescent system. Four additional TTBS washes and a 15 min Tris-Buffered Saline (TBS) (10 mM Tris-HCl pH 7.5, 0.5 MNaCl) wash were done. Detection of the proteins was performed as described previously by chemiluminescent (see section 2.4.6).
Membrane stripping: membrane was washed twice for 5-10 minand incubated30 min at 50ºC in stripping buffer (100 mM β-mercaptoethanol, 2% SDS, 62,5 mM Tris HCl pH 6,7, dH_2O). Then membrane was washed again and ready for new blocking. One membrane can be only subjected to stripping twice for reliable results.

2.5. Affinity chromatography purifications

When needed, stroma samples were concentrated to reduce the sample volume. Concentration wasperformed using Amicon Ultra-4 kit (Millipore), following manufacturer's instructions.

2.5.1. Rubisco elimination from stroma samples

Specific removal of Rubisco protein from stroma samples was achieved by sample fractionation using SepproRubisco Spin Columns, following manufacturer's protocol (SepproRubisco Spin Columns, Sigma-Aldrich). Briefly, samples were passed through a column containing avian (IgY) antigen specifically designed to bind Rubisco. The non-binding sample was collected.

2.5.2. Isolation of glycoproteins containing high mannose type N-glycans

- *Concanavalin A columns (ConA columns):* Isolated stroma fractions, or Rubisco-free stroma samples, were passed through ConA columns (Glycoprotein Isolation Kit, ConA, Pierce) in order to specifically enrich the samples inglycoproteins containing high mannose N-glycans linked to the polypeptide chain. Experiments were performed following manufacturer's instructions. Briefly, the samples wereincubated in columns containing ConA immobilized on an inert matrix, glycoproteins bound to the column were eluted using a buffer containing 200 mM alpha-mannoside which competes with N-glycoproteins for ConA binding, releasing them from the column. In some cases, eluted proteins were dialysed overnight in a Slide-A-Lyzer Dyalisis Cassette (Pierce) with 3500 Da cut off against 20 mM Tris-HCl buffer pH 7.5 to remove the mannoside.

- *Concanavalin A magnetic beads:*Some stroma/chloroplast samples were enriched in N-glycoproteins by purifying with ConA, using a different kit (BruckerDaltonics Inc.), based on the same principle, following manufacturer's instructions.10 µl of sample incubated with 20 µl ConA beads were subjected to several washing steps, and elutedobtaining a N-glycoprotein enriched fraction that was stored for further analysis.

2.5.3. Normalization of minor proteins

In order to reduce high-abundance proteins and concentrate low-abundance proteinsin the chloroplast samples, ProteoMiner Protein Enrichment kit (BioRad), was used following manufacturer's instructions. This method basis relays on the fact that each bead features a different hexapeptide ligand with affinity for specific proteins in a sample. Samples are applied to

the beads, allowing proteins to bind to their specific ligands. Proteins in excess are washed away, and those proteins bound to the beads are eventually eluted, allowing further downstream analysis. Briefly, 200-1200 µL chloroplast samples, containing 50 mg/mL proteins, were incubated 2h at room temperature in rotational shaker at 15 rpm in acolumn containing beads bound tohigh diverse combinatorial peptide ligands, thus reducing the dynamic range of protein concentration(diluting major proteins and concentrating minor proteins), normalizing the amount of each protein present in the sample. The columns were then washed several times in phosphate buffered saline (PBS) (150 mM NaCl, 10 mM NaH_2PO_4, pH 7,4) and then bound proteins were eluted (8 M Urea, 2% CHAPS, 5% acetic acid).

2.5.4. Immunoprecipitation with HA agarose beads

Immunoprecipitation of HA-tagged CAH1 was performed using anti-HA Agarose Conjugate beads (Clone HA-7, Sigma), following manufacturer's instructions. In a few words, 40 µL of anti-HA Agarose Conjugate 1:1 suspension was added to protein extract previously diluted to 200µL with PBS and incubated at 4ºC overnight using an orbital shaker at 20 rpm. Then, samples were washed four times with PBS and the pellet containing precipitated and co-precipitated proteins wastreated with denaturing protein loading buffer (see section 2.4.5) and prepared for further analysis by SDS-PAGE or kept in KEB buffer for endoglycosidase analyses (see below).

2.5.5. N-glycoprotein purification and analysis by LC-MS/MS

All analyses in this section were performed in collaboration with Prof. Michael Hippler's group, in IBBP of Münster University (Germany).

- Glycopeptide enrichment and PNGase treatment by Filter Aided Sample Preparation (N-Glyco-FASP)

Samples were prepared for LC-MS/MS analysis using Filter Aided Sample Preparation, (FASP), allowing minimal handling of samples by performing all reactions in centrifugal filter devices (0.5 ml capacity, 30 kDa MWCO, Amicon Ultra, Millipore) as described inWisniewski et al., 2009 with the following modifications: 200-300 µg of total chloroplast or stroma preparations were subjected to carbamidomethylation of cysteines by adition of 50 mM IAA (iodacetamide) in UA buffer (8 M Urea, 10 mM HEPES, pH 6,5) and 20 min incubation in darkness. Subsequently, tryptic digestionwas performed by adding trypsin (sequencing grade modified, Promega) in a 1:100 ratio enzyme-protein in 50mM ABC buffer (NH_5CO_3 in dH_2O). Trypsin reaction was stopped by addition of 2% formic acid (TFA) solution, and pellet was dried by speed vacuum centrifugation for 3-4 hours. Further enrichment in N-glycopeptides was performed according to

the N-Glyco-FASP protocol (Zielinska *et al.*, 2010) with the following modifications: FASP pellet was reconstituted in 150 µL of TBS-Ca/Mn buffer and mixed with 150 µl of agarose-bound Con A (50 % slurry, Vector Laboratories Inc.), and incubated at room temperature overnight. Then, peptide deglycosylation was carried out using PNGase F (New England Biolabs) or PNGase A (Calbiochem) for 5 h at 37ºC. PNGase F is a N-glycosydase which hydrolyzes nearly all types of N-glycan chains (except those carrying α(1,3)fucose linked to the asparagine-linked N-acetylglucosamine, see Fig. M1). PNGase A was also used (its activity is not affected by the fucose residue mentioned).In some cases,the reactions were performed in presence of ^{18}O-labelled water ($dH_2{}^{18}O$)) in order to indisputably determine whether every aspartic acid residue was already in the polypeptide sequence prior to endoglycosydase digestion, or alternatively it was produced by asparagine deamidation produced by N-glycan loss in the deglycosylation treatment. Then peptides were eluted and samples were dried in a vacuum centrifuge and stored at -20 °C or -80ºC.

- In the case of N-glycoprotein enrichment by *in source* fragmentation analyses, FASP obtained samples were incubated with 200 µl of lectin binding buffer (500 mM NaCl, 1 mM $CaCl_2$, 1 mM $MnCl_2$ in 20 mM Tris-HCl, pH 7.6). Then 200 µL of agarose-bound Con A were added and incubated at room temperature overnight with shaking (900 rpm). Unbound proteins were removed by washing three times with 300 µl NaCl-free binding buffer followed bycentrifugation at 10,000 *g* for 5 min. Glycoproteins were eluted by addition of 150 µl of 0.5 M α-methyl D-mannopyranoside (αMM) in NaCl-free binding buffer with20 min incubation at room temperature with shaking (900 rpm). After centrifugation (10,000 *g*, 5 min) the elution step was repeated once and supernatants were pooled and transferred to a centrifugal filter unit. The samples were centrifuged at 14.000 *g* for 15 min and α-MM was removed by three successive washes with 200 µL of 50 mM ABC buffer. Afterwards, glycoproteins were digested by addition of 2 µg of trypsin in 40µL ABC buffer and overnight incubation at 37°C. Peptides were eluted by centrifugation (14,000 *g*, 10 min). Finally, peptides were dried in a vacuum centrifuge and stored at -20 °C or -80ºC.

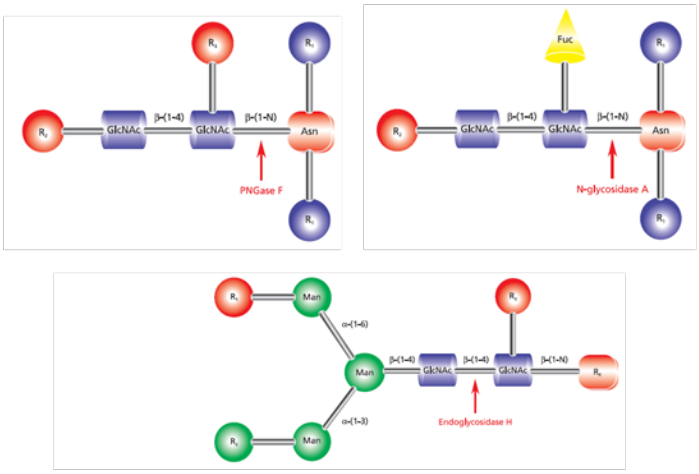

Figure M1: Deglycosilation enzymes. R_1, Oligomannose; R_2, H or mono-or oligosaccharide at the C_2 or C_4 position; R_3, H or α(1,6)fucose; R_4, Asn or Dolichol pyorophospahte. *From "Chemical and enzymatic glycosilation strategies"* Sigma - Aldrich

2.6. Protein Identification by Mass Spectrometry

2.6.1. Matrix-assisted laser desorption/ionization- Time of flight-MS (MALDI-TOF)-MS

These analyses were carried out in the Proteomics Facility of UCM-PCM (Madrid). Spots containing selected proteins (from 1D or 2D gels) were subjected, first to dehydration with acetonitrile (AcN) followed by reduction with 10 mMDTT in 25 mM ammonium bicarbonate (NH_4HCO_3), and alkylation with 55mM Iodoacetamide (10 mg/mL) in 25 mM NH_4HCO_3. Then samples were washed and dried in a vacuum centrifuge. Obtained pellets were digested with trypsin (12.5 ng/ml trypsin solution in 25 mM NH_4HCO_3) and incubated overnight in 25 mM NH_4HCO_3. Then samples were centrifuged and supernatant was taken for peptide extraction with 50% AcN/0.5% TFA solution and sonicated several times. Supernatants were dried in vacuum centrifuged and dissolved in 50% AcN/0,1% TFA solution for analysis with MALDI-TOF Mass Spectrometer Voyager-DE™ STR Biospectrometry Workstation (Applied Biosystems).

2.6.2. Liquid chromatography couplet to Tandem Mass Spectrometry (LC-MS/MS)

All LC-MS/MS analyses were performed in Prof. M. Hippler's laboratory in Münster University (Germany). Chromatographic separation of glycopeptides was performed on an Ultimate 3000 Nanoflow HPLC system (Dionex), which was coupled via a nanospray source to an LTQ Orbitrap XL mass spectrometer (Thermo Scientific). LC-MS mass spectrometer was operated by technical experts in MS of IBBP at Munster University as described in Mathieu-Rivet *et al.,* 2013.

2.6.3. Glycopeptide identification

For peptide identification, three different algorithms were used: X! Tandem CYCLONE (Craig and Beavis, 2004) and OMSSA (Geer *et al.,* 2004), both incorporated into the proteomics data processing pipeline Proteomatic software (Specht *et al.,* 2011), and SEQUEST (Eng *et al.,* 1994) with its own software and operated by technicians at IBBP (Munster) laboratory. MS spectra from the analysed samples were matched against a target-decoy database composed of TAIR10. This database was supplemented with the protein sequence of Con A. Decoy protein sequences were generated by randomly shuffling tryptic peptides while retaining the redundancy of non-proteotypic peptides. The maximum number of missed cleavages allowed was 2. Mass accuracy was set to 5 ppm for MS1 precursor ions and 0.5 Da for product ions. The following modifications were used for all X!Tandem/OMSSA/SEQUEST analyses: Carbamidomethylation of cysteine(static), oxidation of methionine (variable) and deamidation of asparagine (variable). Peptide identifications were statistically validated using Qvality 2.02, (Käll *et al.,* 2008), setting a

q-value threshold of 0.01. Additional variable modifications used for the identification of $dH_2{}^{18}O$-labelled peptides, in this case analysis were performed only with OMSSA and X!Tandem. In the case of X!Tandem, analyses were performed several times on spectra files each time with a slightly modified set of glycosylation related variable modifications, in order to avoid false positives identifications. For that, spectra files were analysed four times by X!Tandem, each time using a different set of variable modifications (Angel *et al.*, 2008): (1) Deamidation of asparagine in $dH_2{}^{18}O$ -labelled water (+2.9883 Da) and single incorporation of 18O at the peptide C-terminus (+2.0043 Da); (2) Deamidation of asparagine in $dH_2{}^{18}O$ -labelled water and double incorporation of $dH_2{}^{18}O$ at the peptide C-terminus (+4.0085 Da); (3) Deamidation of asparagine in dH_2O water (+0.9840 Da) and single incorporation of $dH_2{}^{18}O$ at the peptide C-terminus (+2.0043 Da). (4) Deamidation of asparagine in dH_2O -labelled water (+0.9840 Da) and double incorporation of $dH_2{}^{18}O$ at the peptide C-terminus (+4.0085 Da). All results were combined and conflicting peptide-spectrum matches were filtered on the basis of e-values. If e-values differed by two orders of a magnitude or more, the peptide-spectrum match with the lower score was retained. Otherwise, peptide-spectrum matches were regarded as ambiguous and all corresponding identifications were discarded.However, peptide glycosylation considered as ambiguous were not discarded automatically but validated by manual inspection of fragmentation spectra.

2.7. Biochemical techniques and miscellanea

2.7.1. Chlorophyll measurement

Chlorophyll concentration was measured in total leaf extracts and purified chloroplasts fractions, as described by Porra et al., (1989). An aliquot of the samples was diluted with methanol, mixed and centrifuged 3 min at 14000*g*. The supernatant was placed in a spectrophotometer (MultiSpe-1501, Shimadzu), and chlorophyll *a* (665 nm) and chlorophyll *b* (650 nm) were measured. Total chlorophyll content was calculated following the following formula:

$$[\text{Chlorophyll}]_T = 22{,}1 \cdot \text{Absorbance 650nm (Chlor.}b\text{)} + 2{,}71 \cdot \text{Absorbance 665nm (Chlor.}a\text{)}$$

2.7.2. Calcofluor fluorescence measurement in plate reader

Protoplasts samples placed in a 96-well platewere stained with Calcofluor white M2R (Sigma) to at final concentration of 0,001%. Fluorescence at 350 nm corresponding to calcofluor maximum was measured in a plate reader (BioTek) at different time lapses.

2.7.3. HA tagged CAH1 activity assay

Membrane-inlet mass spectrometry (MIMS) measurements of CAH1 activity were performed in a collaborative work with Prof. G. Samuelsson group in UPSC (Sweden). Measurements were carried out with HC from *Arabidopsis*suspension culture cells stably transformed with HC andimmunoprecipitated as described in Burén et al., 2011, in order to avoid interferences with other CAs isoforms. As control samples, immunoprecipitated proteins from culture cells not expressing the HC and/or buffer were used. HC was immunoprecipitated from culture cells treatedwith tunicamycin for 24 h (10 µg/mL, final concentration). The MIMS measurements of CA activity of all samples wereperformed as described by Clausen et al., 2005, by monitoring the change in $^{12}C^{18}O^{18}O$ concentration (m/z =48) as a function of time after the injection of 15-mL air-saturated $H_2{}^{18}O$ (97% initial; 2.4% finalenrichment). An isotope ratio mass spectrometer (ThermoFinnigan^{Plus} XP) was connected via a cooling trap (dry ice+EtOH) to a home-built membrane-inlet cell similar to thatdescribed by Messinger et al., 1995, but with a largervolume (600 µl). The sample in the cell was separated from thehigh vacuum (3.10^{-8} bar) of the mass spectrometer via a 150 µmthickmetallic mesh silicon membrane (Franatech GmbH) seamlessly resting on a porous Teflon support (Small Parts Inc.). The reaction mixture was kept at 20ºC and stirredconstantly during measurements with a magnetic stir bar. $CO_2$was detected online as the non-labelled ($^{12}C^{16}O^{16}O$), singly labeled ($^{12}C^{16}O^{18}O$), and doubly-labelled ($^{12}C^{18}O^{18}O$) species atm/z = 44, 46 and 48, respectively.

2.7.4. Endoglicosidase-H digestion

N-glycoproteins contained in total leaf extracts, chloroplasts, stroma, protoplasts and other purified eluates were digested with Endoglycosidase H (Endo H) enzyme (New England Biolabs). Endo H is purified from a recombinant *E.Coli* strain, and hydrolyzes the bond between the two proximal GlcNAc moieties in the core N-glycans attached to proteins. Specifically cleaves asparagine-linked mannose rich oligosaccharides (see Fig. M1). Digestion was performing attending manufacturer's indications: Reaction was prepared adding 1-20 µg of N-glycoproteins, 1µL of glycoprotein denaturing buffer 10X, and distilled water till a final volume of 10 µL. Mixture was incubated 10 min at 100ºC. Afterwards, it was added to the reaction 1 µL of G5 reaction buffer 10X, 1µL of PIC 25X, 1µL of PNGase and distilled water till 20 µL of final volume. Reaction was incubated 24h at 37ºC. Once digested, samples were prepared for blotting analysis by adding loading buffer and heating them10 min at 95ºC.

2.8. Microscopy techniques

2.8.1. Fluorescence microscopy

Isolated protoplasts (see section 2.1.4) were stained Calcofluor white MR2 (Sigma) with final dilution of 0,001% and observed with fluorescence microscopy Nikon Eclipse 80i, selecting the DAPI light filter (340-380 nm), as described in Kwon *et al.,* 2005, and visualized with Leica camer and Leica Application Suite v 3.7.0 software (Leica Mycrosystems, GmbH).

2.8.2. Cryoultramicrotomy for immunolabeling of *Arabidopsis thaliana* leaves

Half *Arabidopsis* leaf is placed in a glass plate covered with dental wax on top of a drop of fixative solution (4% paraformaldehide in potassium phosphate buffer (PB) 0,1M (80 mM K_2HPO_4, 20 mM KH_2PO_4, dH_2O, pH 7,4)). Subsequently leaveswere cut in $1mm^2$ squares with a blade and then subjected to vacuum from 5 h to overnight, in order to submerge the material into the fixative solution. Afterwards, samples were transferred to eppendorf tubes and changed to a 2% paraformaldehide fixative solution. Samples are conserved at 4ºC for further processing.
Preparation and visualization of samples for TEM was performed in the Centre Científic I Tecnològic de Universitat de Barcelona. Samples were cryofixed by Leica EMAFS cryofixation systems, and ultrathin sections (50 - 70 nm) were obtained from cryofixed leaf tissue, by using Leica Ultracut UCT ultramicrotome with Leica EMFCS cryoultramicrotomy unit at temperature between -120ºC and -100ºC, and mounted in nickel grids for immunogold assays.

2.8.3. Immunocitochemistry

This experimental procedure was performed the Centre Científic I Tecnològic de Universitat de Barcelona (PCB) Service. Cryosections obtained from fixated leaves of *Arabidopsis* wild type and FT-myc/HAexpressing lines were loaded in proper grids covered by sucrose, and subjected to immunocitochemistry (immunogold), in order to localize labeled proteins within the plant cell.Grids with cryosections were incubated for 30 min in 2% PHEM gelatin at 37ºC, in order to remove sucrose cover. Grids were washed with 0,15 M glycine in PHEM buffer (60 mM PIPES, 25 mM Hepes, 10 mM EGTA, 2 mM MgCl2; pH 6.9) for removing aldehyde groups from fixation. Washes of 10 min in 10% FBS in PHEM were made for blocking the grids. Then washes were performed in 1% FBS in PHEM during 2 min. Then, samples were incubated with the desired dilutions of primary antibodies (HA, Covance or myc Santa Cruz) in PHEM with 1% FBS, 1h. Afterwards, excess of primary antibody was removed by 8 washes of 2 min in 0,2% FBS, and a last wash in 1% FBS 2 min. Then samples were incubated in desired dilution of gold

labeled secondary antibody (goat anti-mouse IgG-conjugated 10 nm gold particles, Sigma, St.Louis, MO, USA, for samples incubated with HA, or anti-rabbit IgG-conjugated 10 nm gold particles, Sigma for samples incubated withmyc)for 30 min. All steps from here on were performed in darkness. After incubation, 3 quick washes were performed in 0,1 M PHEM and 7 washes of 2 min were performed in 0,1 M PHEM. Then, grids were incubated in glutaraldehyde in 1% PHEM for 5 min. Subsequently there were performed 10 washes of 1 min each with dH_2O, andgrids were incubated in 2% uranyle acetate pH 7 for 5 min. 3 washes in dH_2O and 10 min of incubation in uranyle acetate – metyl cellulose. Grids were dried 30 min, and were ready for TEM observation.

2.8.4. Transmission electron microscopy (TEM)

Visualization of samples by TEM was performed inthe Centre Científic I Tecnològic de Universitat de Barcelona. Microscope analysis was performed in a JEOL JEM 1010 transmission electron microscope (Tokio, Japan), set to 80 KV. Images were captured using a BioScan camera (Gatan, Pleasanton, CA, USA).

2.9. Bioinformatic techniques

- Primer design was performed using the software at the NCBI site, Primer-Blast (Ye *et al.*, 2012).
- Sequences analysis to check proper insertions in the cloned vectors was performed using UGENE software (Okonechnikov *et al.*, 2012).
- Sequence alignments were performed using Clustal Omega software (Larkin *et al.*, 2007)
- General search of the proteins identified by mass spectrometry experiments were performed in the following online databases: SUBA3 (Tanz *et al.*, 2013), TAIR (Lamesch *et al.*, 2012), PPDB (Sun *et al.*, 2009).
- N-glycosylation consensus sequence and signal peptide predictors: NetNGlyc (Gupta el al., 2004)and Target P (Emanuelsson *et al.*, 2000).
- Pixel intensity and image densitometric analysis for gene expression and protein immunodetection experiments: Obtained images of genetic samples from RT-PCR experiments, capturing intensity of emitted UV fluorescence, were quantified using Image J software (Schnaider *et al.*, 2012). For that, equivalent regions of interests (ROIs) were selected, measuring their intensity. Data of housekeeping gene (*adp-rf*) was taken as reference, and intensity ratios KOR1/ADP-RF were calculated and subjected to statistical analysis.

Densitometric analysis of proteomic samples was carried out with the ImageLab software (BioRad), in an analogous modeas described before: selecting areas of interest in non saturated images and measuring their pixel intensity. In all cases, intensity of loadings controls were used as references, and normalization of measured samples were performed by calculating the ratio of intensity of band sample (HA signal for IP assay/intensity of band of loading control –BiP-). In the case of Endo H digestions, the ratio was calculated by measured intensities of band resistant/non-resistant.All data was further processed for statistical analysis.

2.10 Statistical analysis

Statistical analysis of crude data was performed using software R 2.15.0 (R Core Team., 2014). Shapiro Wilk analysis was applied to test the normality of the data. For data with normal distribution, ANOVA was applied in order to test the significance of the relation subject of study. In case of data with no normal distribution (that is, non parametric), Kruskal-Wallis test was applied. Sample size “n”was considered as technical replicates, taking into account in every moment their original biological replicate and, in specific cases, the treatment it was subjected.

VIII: RESULTS

CHAPTER 1: IMPORTANCE OF POST-TRANSLATIONAL MODIFICATIONS FOR PROPER FOLDING OF A CHLOROPLAST-LOCALIZED αCARBONIC ANHYDRASE (CAH1)

The study performed in 2005 (Villarejo *et al.*, 2005) showed for first time a protein, an α carbonic anhydrase, (CAH1, AT3G52720) of *Arabidopsis thaliana* (*Arabidopsis*) becoming N-glycosylated when passing through the endomembrane system, and reaching the chloroplast afterwards. CAH1 is an α carbonic anhydrase that catalyzes the inter conversion of CO_2 and HCO_3^-, an essential process in photosynthetic reactions (Fabre *et al*., 2007). The discovery of this non canonical pathway for nuclear encoded proteins to an organelle, suggested that proteins using this pathway require those post-translational modifications acquired during the pass through ER and Golgi, for proper targeting and function.

In a collaborative work with Prof. Samuelsson´s group in Umeå Plant Science Centre (UPSC) (Umeå, Sweden), the importance of CAH1's PTMs, i.e, N-glycosylations, together with potential disulphide bonds, and the significance of the C terminus sequence, has been evaluated. The study has been focused in the analysis of the protein stability and traffic, by characterizing the influence of named PTMs on protein folding. With that purpose, a hemagglutinin epitope (HA) fused to the N terminus of the CAH1 sequence was used (Fig. 1.1). Although fusions to short epitopes like HA do not usually alter significantly the tertiary structure and biologic activity of the native protein (Zhao *et al.*, 2013), some parameters were checked. First, the chloroplast localization of CAH1 (Villarejo *et al.,* 2005) was confirmed, in this case of the HA-tagged version (named HACAH1 or HC) by immunogold labeling and transmission electron microscopy (TEM) (see Burén *et al.,* 2011). This prompted us to continue the study using this fusion protein as a valid tool. In first term, the N-glycosylation pattern of HC was analyzed, and it was observed that it presented a heterogeneous mix of complex and high mannose glycoforms (see Suppl. Fig.S1.1, from Burén *et al.*, 2011).

It was subsequently decided to generate and characterize a series of mutated HC versions, each one presenting a modification in a key aminoacid, necessary for acquisition of posttranslational modifications like N-glycosylation or disulfide bonds formation. Mutants disrupting the peculiar C terminus of the protein were also created.

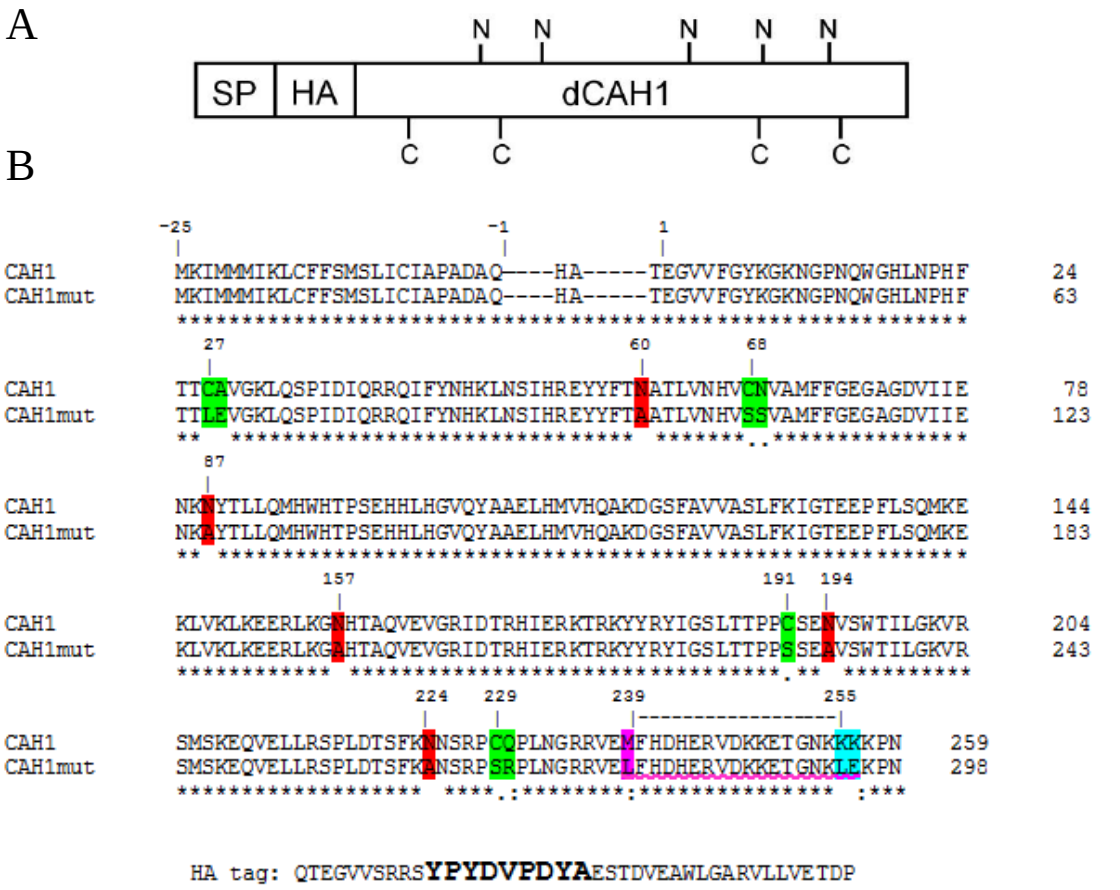

Figure 1.1. Schematic view and aminoacid sequence of the HA tagged CAH1 (HC) and its mutated versions. A) Diagram of the fusion protein. There are pointed the position of the fused HA epitope, the five potential N-glycosylation sites (N) and the four cysteine residues (C), B) Schematic representation of the different point mutations performed in HC: Cysteine residues (green), N-glycosylation sites (red), and C terminus deletion (pink) and mutation (blue). Position and sequence of HA tag is indicated. SP, signal peptide; dCAH1, native protein. Adapted from Burén *et al*., 2011

1.1 *In silico* study and generation of point mutated versions of HC

The mature CAH1 protein has five potential N-glycosylation sites, and contains four cysteine residues potentially forming two disulphide bonds (Fig 1.1, 1.2, and Burén *et al.*, 2011). By site directed mutagenesis, mutant variants of HC, cloned in a modified pPE1000 vector (pHACAH1) (see section 2.2.7 of Material and Methods) were created. These modifications consist in disruption of N-glycosylation sites, cysteine residues potentially involved in disulphide bond

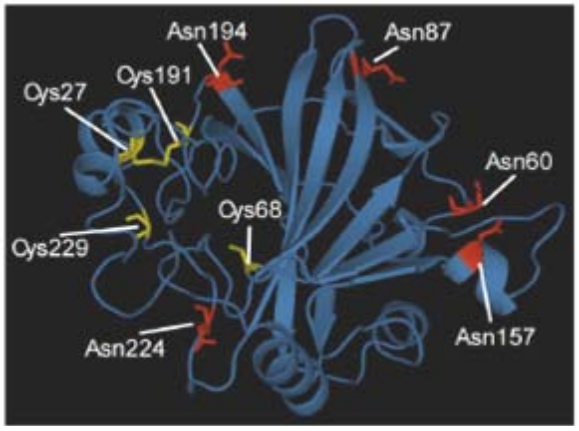

Figure 1.2. Prediction of the 3D structure of CAH1 using SWISS-MODEL (Kiefer *et al.,* 2009) and visualization using Swiss-PdbViewer (Guex *et al.,* 1997). Figure shows the five potential N-glycosylation sites (red) and the four cysteine residues of the mature CAH1 protein (yellow). From Burén *et al*., 2011

Table 1.1. Constructions used to transfect protoplasts from cell cultures and mesophyll tissue of *Arabidopsis thaliana*, and agroinfiltration *of Nicotiana benthamiana* leaves

Clone	Mutation	Mutated sites	Description
HC	-	-	HA-tagged wt CAH1
C1	CA27LE	Cysteine 1	
C2	CN68SS	Cysteine 2	
C3	C191S	Cysteine 3	
C4	CQ229SR	Cysteine 4	
C1+C3		Cysteine 1 and 3	= C1+C3
N1	NAT60AAT	Glycosylation site 1	
N2	NYT87AYT	Glycosylation site 2	
N3	NHT157AHT	Glycosylation site 3	
N4	NVS194AVS	Glycosylation site 4	
N5	NNS224ANS	Glycosylation site 5	
N1+N2		Glycosylation sites 1 and 2	
N3+ N5		Glycosylation sites 3 and 5	
N3 +N4		Glycosylation sites 3and 4	
N4 + N5		Glycosylation sites 4 and 5	
N3 + N4 + N5		Glycosylation sites 3, 4 and 5	
Q1		Glycosylation site 2, 3, 4 and 5	= N2+N3+N4+N5
Q2		Glycosylation site 1, 3, 4 and 5	= N1+N3+N4+N5
Q3		Glycosylation site 1, 2, 4 and 5	= N1+N2+N4+N5
Q4		Glycosylation site 1, 2, 3 and 5	= N1+N2+N3+N5
Q5		Glycosylation site 1, 2, 3 and 4	= N1+N2+N3+N4
NG		All glycosylation sites	= N1+N2+N3+N4+N5
Cdel	Deletion from M>L239 to K>E256	16 aminoacids deletion of C terminus	
Cmut	KK255LE	Lysine substitution	

formation, and alteration of C terminus, as described in Fig. 1.1 B and Table 1.1, following the protocol described in section 2.2.7 of Material and Methods. Briefly, high fidelity PCR was performed using pHACAH1 as DNA template, with primers introducing the point mutations indicated in Fig. 1.1B. In the case of N-glycosylation mutants, phosphorylated primers were used, and the whole plasmid and insert were amplified and further ligated by the phosphorylated ends. In the case of cysteines and C terminus mutations, PCR amplified fragments were digested and cloned into a pPE1000 modified vector. The obtained constructions were transiently expressed in protoplasts from both, *Arabidopsis* cell cultures and leaf mesophyll tissue, and in *Nicotiana benthamiana* (*Nicotiana*) intact leaves. Protein extracts from these biological samples transiently expressing the constructions were subsequently subjected to biochemical analysis, and tested in SDS-PAGE gels.

1.2 N-glycosylation is required for correct CAH1 folding

1.2.1 HC presents four or five N-glycosylated sites

Initially, all N-glycosylation constructions (N1-N5, double and triple combinations, and Q1-Q5, see Table 1.1) were transiently expressed in *Arabidopsis* cell culture protoplasts. Protein extracts were analyzed by SDS-PAGE gels followed by immunoblot detection using anti-HA antibodies. First, two prominent HC glycoforms were detected (Fig. 1.3), in accordance to the previously mentioned heterogeneous pattern of glycoforms. Also the apparently occurrence of prominent bands in N1, N2, N3 and N5 mutated versions, similar to HC pattern, was observed. Suprisingly, N4 mutation produced only a single band, of the size of the smaller prominent band of HC.. However, the molecular mass of the bands corresponding to the different N-glycosylation sites mutants was lower than that of HC, reflected as a higher mobility of these proteins (Fig. 1.3A). This is consistent with the lack of glycan(s) residue(s) which should be occupying their respective sites. Differences in molecular mass are proportional to the lack of one N-glycan (single mutants, N1 to N5), two glycans (double mutants, N_x+N_x), three glycans (triple mutant), four glycans (quadruple mutants, Q) and absence of all N-glycans (NG construction). Q mutants, which only harbor one N-glycan, consistently presented a slightly higher molecular mass than NG. Interestingly, the Q4 mutant, mutated in N1, N2, N3 and N5 sites, migrates as two bands, one similar to the other quadruple mutants harboring only one N-glycan residue, and other

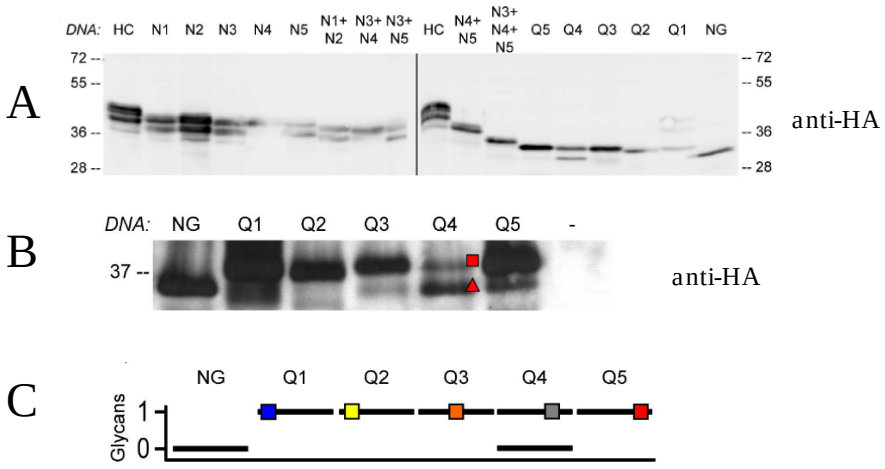

Figure 1.3: HA tagged CAH1 (HC) harbours four or five N-glycans. A) Protein extracts from *Arabidopsis* protoplasts from culture cells, transiently expressing HC and its N-glycosylation mutated variants. Migration patterns of the different mutants are showed. B) Protein extracts form *Arabidopsis* mesophyll cells expressing mutant N-glycosilation sites variants from HC (NG and Q1-Q5). Mutation of four of the five N-glycosylation sites (Q1-Q5) leads to polypeptides that migrates slightly slower than NG. Q4 migrates as two bands, one of the same size as the others Q mutants (square), and other like NG (triangle). Non transformed cells are used as negative control, C) Schematic view of B. See Table 1.1 for abbreviations and Figure 1.1 for localization of N-glycosylation sites. From Burén *et al*., 2011

considerably more intense, with molecular mass corresponding to the NG mutant, suggesting that this N-glycosylation site is not always occupied (Fig. 1.3B and C). This is also in accordance with the previously mentioned N4 pattern, in which is only observed a single band whose size correspond to a ttachment of four N-glycans (Fig. 1.3A). Taken together, these results suggest that, in HC, N-glycosylation site 4 is occasionally *empty*.

1.2.2 Confirmation of non glycosylation of "NG" construction

Once the occupancy of N-glycosylation sites of the different versions of HC was checked, the confirmation of the non glycosylated state of NG construction was carried out. For that, in a collaborative work with a group in UPSC, protein extracts from protoplasts of *Arabidopsis* cell cultures, transiently expressing the constructions, were treated with Tunicamycin, which inhibits N-glycosylation of proteins in the ER (Park et al., 2004). NG mobility shift was compared to that produced on HC treated with Tunicamycin, using non treated HC as a control. Treated samples yielded a band with a molecular mass equivalent to that produced by the disruption of all CAH1 N-glycosylation sites (NG) (Fig. 1.4). This mobility shift is consistent with the mass difference previously observed in deglycosylated native protein *in planta* (Villarejo *et al.*, 2005).

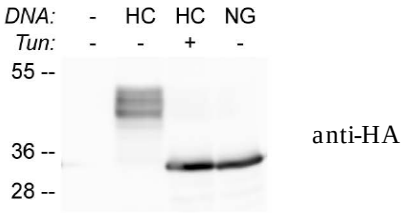

Figure 1.4: Confirmation of lack of N-glycans in NG mutant. Protein extracts of *Arabidopsis* mesophyll cells, showing heterogeneous migration pattern of HC in normal conditions, and subjected to Tunicamicyn (Tun) treatment, in comparison with pattern of non glycosylated HC (NG). From Burén *et al.,* 2011

Transient expression of the mutant variants is tested in two different model systems

Folding is a key feature for proper protein function. With the aim of studying this item, HC and the mutant variants NG were transiently expressed in protoplasts from *Arabidopsis* mesophyll tissue, and in leaves of *Nicotiana.* Subsequently, the nature of N-glycans attached to the protein, and the degree of folding of the resulting proteins was analyzed.

Until this point of the present work, constructions were transiently expressed in *Arabidopsis* cell culture protoplasts. Cell cultures protoplasts (plant cells lacking the cell wall) are a very extended and useful tool to study intracellular trafficking, especially when transient expression of recombinant proteins labeled with fluorescent markers are tested using microscopy techniques (Miao *et al.*, 2007; Bargmann *et al.*, 2010). However, in some situations, although protoplasts are considered as a generally reliable system for quick protein analysis, other studies have reported different responses and variability in their expression analysis due to the alterations produced by cell wall removal that can affect different cellular functions (Lee *et al.*, 2006). Therefore, it has been argued that they cannot be used as the only model to study a trafficking process.

Alternatively, transient expressions systems using *Agrobacterium tumefaciens* (*Agrobacterium*) mediated transformation of plant leaves are also extended, although microscopy studies in this system occasionally are more problematic. However, agroinfiltration in *Nicotiana* leaves have the advantage of being a more realistic system, where cells are placed in their tissues receiving the different signals produced by the surrounding cells. In addition, *Nicotiana* has a short life cycle, and possess relatively large and easy infiltrable leaves that produce recombinant proteins at high levels, providing sufficient amount of biological material for further experiments (Ma *et al.*, 2012).

In order to avoid possible artifacts due to the model system used for the studies, it was decided to use both of them: in first term, by performing a test in *Arabidopsis* protoplasts isolated from mesophyll tissue, instead of cell culture, to get an idea of the protein expression and accumulation rates, and concominantly see whether this isolation method yields better results than protoplasts from cell cultures. Secondly, by transforming *Nicotiana* leaves to further analyze/confirm the parameters studied.

1.2.3 N-glycosylation is important for proper protein folding and trafficking

Traffic of mutated versions of HC

The use of endoglycosydases is an approach traditionally used to determine the pass of glycoproteins through different organelles within the endomembrane system, since it is generally accepted that, the presence of high mannose N-glycans indicates proteins remain or are retained in the ER (Miyazono *et al.*, 1992; Chen *et al.*, 2002). In order to have some insights about traffic of CAH1 mutant versions, an approach combining use of endoglycosydases and co-immunoprecipitation with BiP chaperone was carried out (see following sections). On one side, proteins are subjected to digestion of Endo H, an enzyme that can deglycosylate proteins that have not been subjected to mannosidase II (ManII) action (i.e., proteins containing the so called *high mannose N-glycans.* Fig. I3l). As previously mentioned, Endo H releases the N-glycan residues attached to glycoproteins leaving the core N-acetylglucosamine attached to the Asn of the protein (section 2.7.4 of Material and Methods). Since ManII is located in the cis-Golgi compartment, the glycoproteins that have reached this point should undergo N-glycan processing leading to hybrid or complex N-glycans, and in this case, they would not be deglycosylated by Endo H. HC and the mutated versions were transiently expressed in *Arabidopsis* protoplasts from mesophyll tissue and *Nicotiana* leaves.

- *NG is resistant to Endo H*

In the case of NG, no alterations were produced by Endo H treatment, confirming, from another point of view, the lack of N-glycans attached to the structure (Fig. 1.5A and D). The very low signal obtained in constructions expressed in protoplasts made hardly appreciable protein bands in the Western blot experiments. As could be expected, protoplasts samples failed to show protein bands in further analysis, ie, immunoprecipitation. Therefore, these constructions were cloned into pMDC32, a plant expression vector, which allows *Agrobacterium* mediated transient expression in *Nicotiana* leaves (Curtis *et al.*, 2003, section 2.3.2 of Material and Methods). In this case, results were confirmed *in planta*. First, effective protein expression was checked with *green fluorescence protein* (GFP) cloned in the pMDC32 vector (Fig. 1.5E and F). In all immunoassays, transient expression in *Nicotiana* showed higher protein levels, in every biological replicate, than those observed in protoplasts (Fig. 1.5C and D). These samples clearly confirmed the absence of N-glycan in NG construction, as well as the HC partial digestion, supporting the heterogeneous nature of the N-glycans attached to the structure.

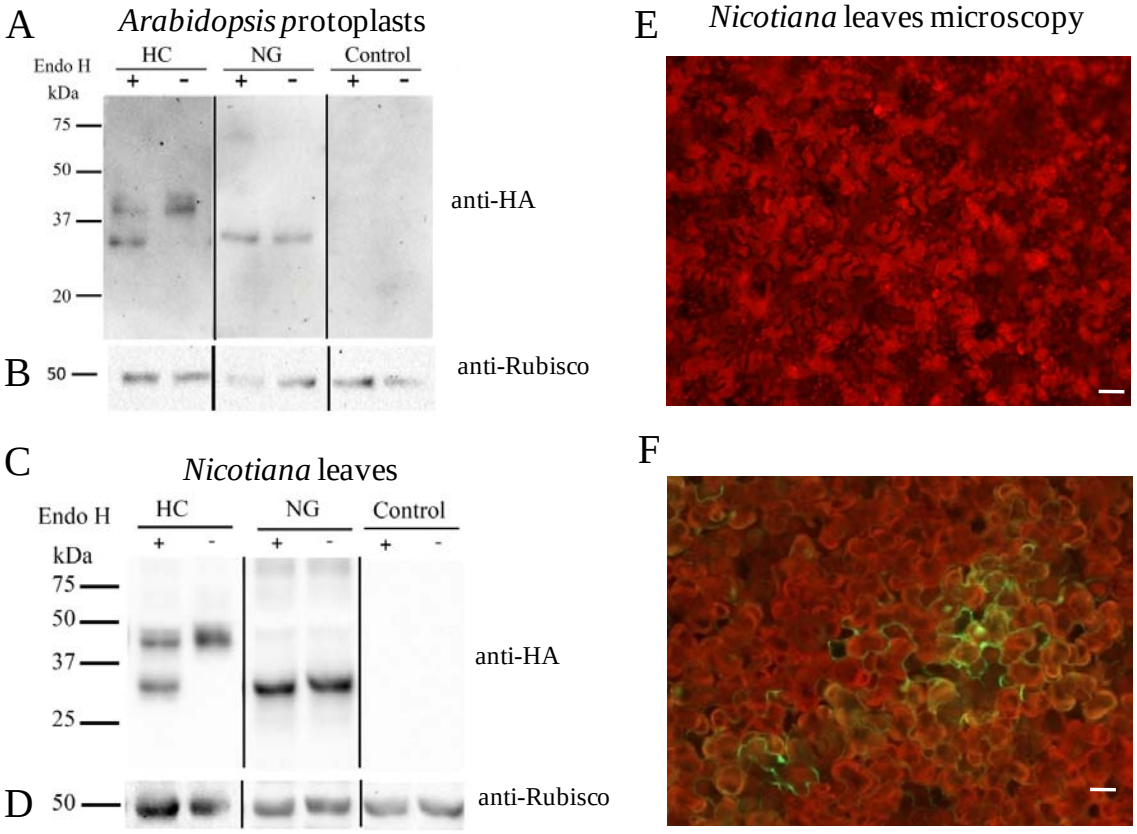

Figure 1.5: Transient expression of NG mutated version of HC in *Arabidopsis* protoplasts and *Nicotiana* leaves. Protein extracts were prepared from isolated protoplasts from mesophyll (A and B) and *Nicotiana* leaves (C and D) expressing HC and NG. Water transfected protoplasts or non transformed *Nicotiana* leaves were used as negative control. Black line indicates different areas of same blotting experiment. A and C show immunoblot with HA antibody in samples subjected to Endo H treatment (+), control reactions are marked with (-). B and D show loading control with Rubisco antibody. E and F) Control of transient expression of constructions in *Nicotiana* leaves expressing GFP observed at optical microscope. E) Samples observed in bright field, F) samples observed with blue light filter. Scale bar 50 µm.

- *N-glycosylation is important for proper folding of protein*

As mentioned before, to test protein folding, co-precipitation assays with an ER resident chaperone was performed. The aim was to test whether protein folding is affected, by assessing the degree of union to chaperone BiP. It is known that N-glycosylation plays a role in protein folding and stability (Parodi 2000; Mitra *et al.*, 2006; Price *et al.*, 2010; Kaushik *et al.*, 2010), and is generally accepted that lack of glycosylation leads to protein oligomerization and aggregation in the ER (Parodi 2000; Wujek *et al.*, 2004; Mitra *et al.*, 2006; Kang *et al.*, 2008; Burén *et al.*, 2011). This fact provokes the triggering of the Unfolded Protein Respose/ER Quality Control System, consisting in chaperone-mediated retention of proteins with aberrant conformation, in order to subsequently transport them out to the cytosol for proteasome degradation (Hong *et al.*, 2008; Li *et al.*, 2009). One of these chaperones is the *binding immunoglobulin protein*, BiP, which binds to misfolded proteins, among other functions (Sung *et al.*, 2001; Hebert and Molinari, 2007).

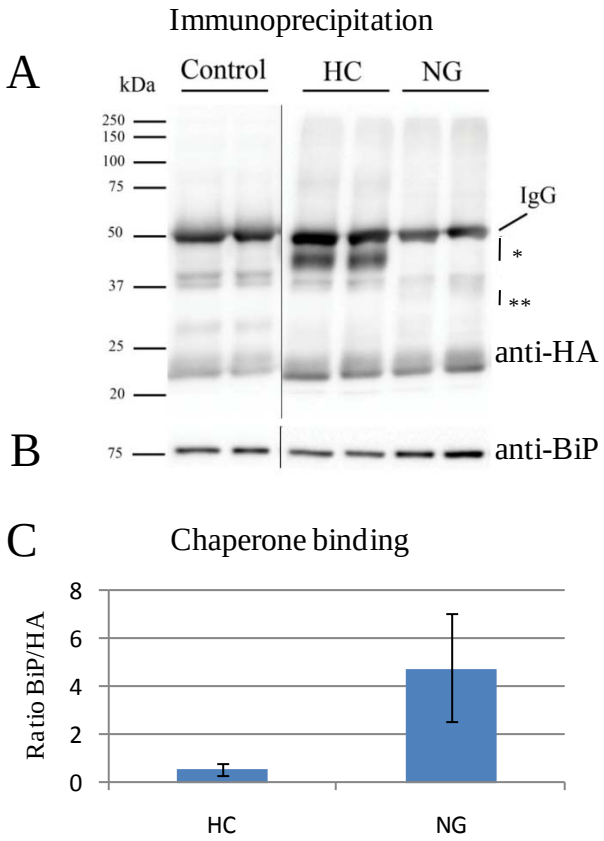

Figure 1.6: N-glycosylation is involved in proper folding of HC. Protein extracts from total leaf extract from *Nicotiana* expressing HC and NG mutant version were immunoprecipitated using HA agarose beads. As negative control, immunoprecipitated total extract from non transformed plant was used. Black line indicates different areas of same blotting experiment. A) Immunoblot with HA antibody. (*) indicates immunoprecipitated HC protein, (**) indicates immunoprecipitated NG protein. B) Immunoblot with BiP antibody. C) Densitometric analysis of immunoblots A and B, estimating the relative amount of BiP that co precipitates with the mutant tagged versions of CAH1. The graph represents the ratio of BiP to HA-tagged protein (mean ± SE, n=24, p<0,001. ANOVA test) IgG, immunoglobuline G.

Interaction between BiP and misfolded proteins can be detected by immunoprecipitating the misfolded protein, which leads to BiP co-precipitation (Forsayeth *et al.*, 1992; Kjaer and Ibáñez, 2003; Burén *et al.*, 2011). For that, protein extracts from transformed *Nicotiana* plants were incubated with anti-HA agarose beads, allowing HA-tagged protein purification/precipitation from the complex sample, and therefore enabling a better separation and further visualization in SDS-PAGE (see section 2.4.5 of Material and Methods). Western blot analysis was performed using anti-HA (Fig. 1.6A) and anti-BiP antibodies (Fig. 1.6B). Subsequently, a pixel

densitometric analysis of such blots was performed (Fig. 1.6C). Relative high amounts of BiP co-precipitated in NG samples when compared to HC, meaning that chaperone binding rate was higher and the ERQC was activated by the non-glycosylated protein. Therefore, results show that complete lack of N-glycosylation affects protein folding, as high amounts of chaperone BiP binds to NG (Fig. 1.6A).

These results, reflecting a clear influence of N-glycosylation on protein folding and trafficking, prompted us to continue investigating the influence of the potential disulphide bonds and the information contained in C terminus of CAH1.

1.3 Intramolecular disulphide bonds are important for folding and ER export of CAH1

There are evidences of disulphide bonds being important for proper folding and stability of the proteins, as well as for function (Tu and Weissman, 2004). In addition, a relationship between disulphide bonds and catalytic activity in other αCAs has been reported (Waheed *et al.*, 1996; Hilvo *et al.*, 2005).

1.3.1 There is a disulphide bond between cysteines C1 and C3

Mature protein CAH1 possess four cysteine residues which could form potentially two disulphide bonds, one between cysteines (C) 68 and 229 (C2 and C4 respectively from here on), and other between residues 27 and 191 (C1 and C3 respectively). Considering the predicted three dimensional model of the protein, CAH1 structure seems to present close proximity between C1 and C3 (Fig. 1.2). In a collaborative work with Prof. Samuelsson's group, single cysteine mutants (C1-C4, Table 1.1) were separated in a SDS-PAGE gel, under reducing and not reducing conditions (i.e., in the presence/absence of β mercaptoethanol (2-ME), a reducing agent that cleaves disulphide bonds. Under non-reducing conditions (i.e., in absence of 2-ME), C1 and C3 mutants presented a differential migration pattern when compared to HC (Figure 1.7A, upper panel). However, under reducing conditions, HC and all cysteine mutants migrated with a similar

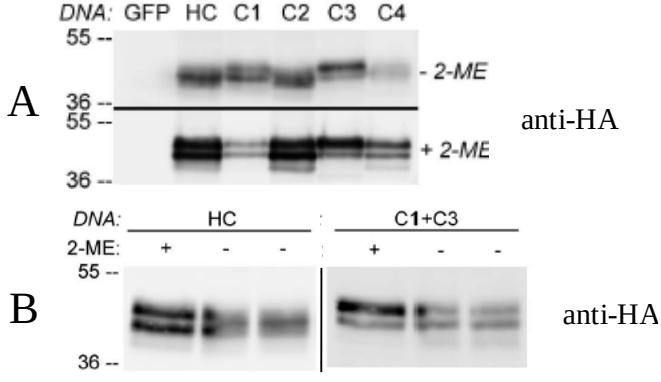

Figure 1.7: Intramolecular disulphide bond in cysteines C1 and C3 of HC. Immunoblot analysis with HA antibodies of protein extracts from *Arabidopsis* cell suspension culture, transiently expressing HC and cysteine mutant variants. There were obtained protein patterns under non reducing conditions (-2ME) and reducing conditions (+2ME). As negative control, protoplasts expressing GFP were used. A) Protein pattern of HC and single cysteine mutants B) Comparison of protein migration pattern between HC (left panel) and double cysteine mutant C1C3 (right panel). From Burén *et al.,* 2011

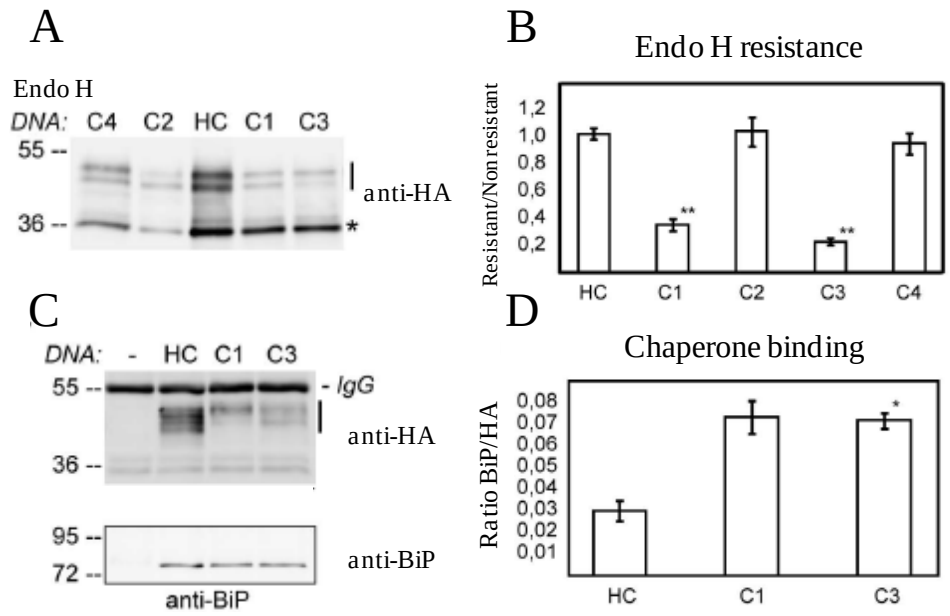

Figure 1.8: Endo H digestion in cysteine mutants support the existence of an Intramolecular disulphide bond in HC. Protein extracts from transfected protoplasts from *Arabidopsis thaliana* culture cells. A) Endo H digestion of HC and single cysteine mutants. Resistant glycoforms are marked by a line, asterisk (*) indicates sensitive glycoforms. B) Densitometric analysis analysis of the relative amounts of Endo H resistant and sensitive glycoforms. The graph represents ratio of resistant/sensitive proteins (mean ±SE, **p<0.01, n=4). C) Immunoprecipitation with HA agarose beads. Level of coprecipitation with chaperone BiP was analyzed with HA (upper panel) and BiP antibodies (lower panel). Immunoprecipitated HC glycoforms are marked by a line. Heavy chain of immunoprecipitation is indicated as *IgG.* D) Densitometric analysis of BiP binding, shown as ratio of BiP to HA tagged CAH1.(mean ± SE, *p<0.05, n=2). From Burén *et al.,* 2011

pattern (Figure 1.7A, lower panel). This was consistent with the existence of a disulphide bond between C1 and C3 residues, as mutated protein seem to be affected in their structure, probably due to the actual lack of disulphide bond. In addition, C1C3 double mutant did not show a mobility shift, being apparently not affected by the addition of 2-ME (Figure 1.7B right panel), as protein pattern was similar to HC, meaning that there is no any union/bond being disrupted by the reducing agent (Figure 1.7B left panel). Conversely, C2 and C4 mutations did not seem to produce any significant effect on CAH1 mobility (Fig 1.7A).

Regarding protein stability, it was postulated that the disulphide bond between C1 and C3 is important for proper folding of CAH1. Therefore, it could be expected that disruption of such bond could lead to the protein accumulation in the ER, and BiP chaperone binding. Co-immunoprecipitation studies of single cysteine mutants were performed, corroborating the existence of a disulphide bond between C1 and C3, as these mutant versions produced enhanced BiP co-precipitation compared to wild type HC (Fig. 1.8). Nevertheless, further studies in *Arabidopsis* protoplasts from mesophyll tissue and *Nicotiana* leaves were performed.

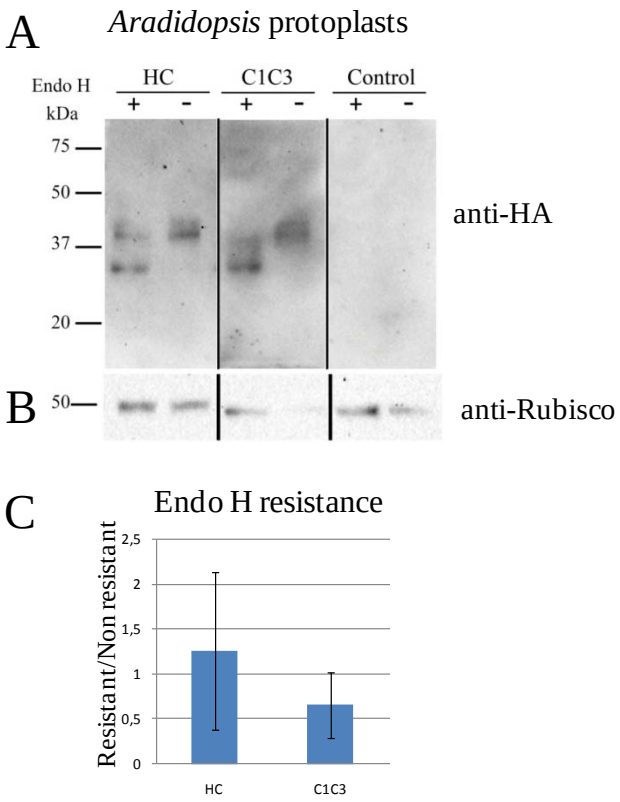

Figure 1.9: Transient expression of C1C3 mutated version of HC in *Arabidopsis* protoplasts. Protein extracts from iprotoplasts from mesophyll tissue expressing HC and C1C3. Water transfected protoplasts were used as negative control. Black line indicates different areas of same blotting experiment. A) Immunoblot with HA antibody. Endo H treatment (+) and control (-). B) Loading control with Rubisco antibody. C) Densitometric analysis of immunoblot A, representing the relative amounts of Endo H resistant and sensitive glycoforms. The graph represents the ratio of Endo H resistant/sensitive glycoprotein (mean ± SE, n=24). No statistical significance, high variability data (AVOVA test).

1.3.2 C1C3 is sensitive to Endo H

For further analysis, double cysteine mutant C1C3 was transiently expressed in *Arabidopsis* protoplasts from mesophyll cells and subjected to Endo H treatment (Fig. 1.9). It could be appreciated that C1C3 double mutant was highly sensitive to the Endo H deglycosylation, meaning that the protein mainly harbors high mannose N-glycans. This fact implies some relationship between the presence of disulphide bond and the type of N-glycosylation acquired. It can be noted that the intensity of the recorded anti-HA signal was very low, probably due to low protein expression. This is consistent with previously obtained results in the collaborative work with Prof. Samuelsson's group, in which systematically lower protein level was perceived in all mutants involving C1 and/or C3 alteration (see also Burén *et al.*, 2011).

Unfortunately, immunoprecipitation of C1C3 double mutant transiently expressed in protoplast from *Arabidopsis* mesophyll tissue, also produced a very low signal in western blot, hardly distinguishable from the membrane background, similar to that of NG expression in mesophyll protoplasts. Indeed, as previously mentioned, this happened in all constructions expressed in protoplasts, therefore these results are not shown. Additionally, C1C3 double mutant construction could not be cloned into pMDC32 vector for plant transformation, impeding expression it in *Nicotiana* leaves, but corroborating the difficulty of handling this construction, from what it could be deduced that disulphide bond must be decisive for proper functioning of the protein.

1.4 C terminus is necessary for correct folding of CAH1

As mentioned in Introduction, it was observed that the C terminus of CAH1 has special characteristics, as is highly hydrophilic residues and lysine enrichment (Villarejo *et al.*, 2005). It was considered it could contain some kind of targeting information of the protein. In order to study its influence in protein folding, two versions with altered C terminus were created.

1.4.1 Cdel shows highly sensitiveness to Endo H, Cmut is partially sensitive to Endo H

For this purpose, two constructions were made: "Cdel", in which a section of the C terminus sequence was deleted, and "Cmut", where the two central lysines of polylysine domain were mutated to LE (Fig. 1.1 and Table 1.1). Constructions were transiently expressed in *Arabidopsis* protoplasts and *Nicotiana* leaves (Fig. 1.10). In both cases were subjected to Endo H digestion as described before.

"Cdel" construction showed high sensitiveness to Endo H treatment, meaning that N-glycans

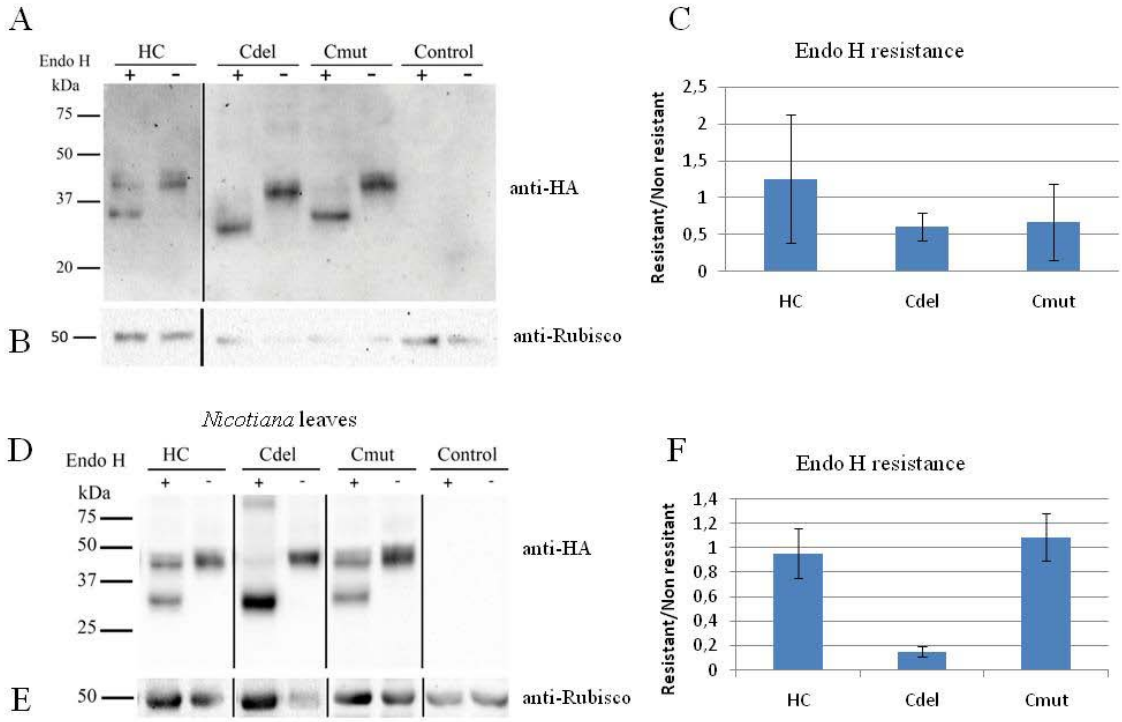

Figure 1.10: Transient expression of C terminus mutated versions of HC in *Arabidopsis* protoplasts and *Nicotiana* leaves. Protein extracts from protoplasts from mesophyll or total leaf extract of *Nicotiana* expressing HC and C terminus mutated versions of HC. Water transfected protoplasts or non transformed *Nicotiana* leaves were used as negative control. Black line indicates different areas of same blotting experiment. A and D) Immunoblots with HA antibody. Endo H treatment (+) and control (-). B and E) Loading control with Rubisco antibody. C and F) Densitometric analysis of immunoblots A and B, in the case of protoplasts, or D and E in the case of *Nicotiana*, representing the relative amounts of Endo H resistant and sensitive glycoforms. Graphs represent the ratio of Endo H resistant/sensitive glycoprotein (mean ± SE, n=24. No statistical significance due high variability data in the case of protoplasts. p<0,001 in the case of *Nicotiana*, ANOVA test.

attached to the structure are mainly high mannose type. This could mean that important information is contained in the C terminus of the protein, as the mutated protein harbors not noticeable complex glycans but high mannose. Interestingly, “Cmut” mutant variant presented the same partial digestion as control HC, clearly appreciable when expressed in *Nicotiana*, suggesting that the change of charge in this part of the sequence do not seem to affect the protein structure, which presents the same or very similar glycosylation pattern as HC. Densitometric analysis was performed subsequently, showing the high sensitivity of Cdel to the enzyme, and moderate levels

of sensitivity, similar to HC, in Cmut construction. It can be observed, as in the case on NG, that transient expression in *Nicotiana* leaves provide far more appreciable signal in Western blots. Indeed, statistical analysis of data from densitometric analysis was significant in the case of *Nicotiana*, while high variability hamper significant results in analysis of data from protoplast system.

1.4.2 Cdel binds moderately to chaperone BiP, while Cmut shows similar behavior than HC

As described for NG construction, co-immunoprecipitation of chaperone BiP was tested in C terminus mutated versions of CAH1. As it can be observed in Fig. 1.11A and B, Cdel showed moderate affinity to BiP, while Cmut presented a similar behavior than HC, being consistent with the observed sensitivity to Endo H. Densitometric analysis measuring the degree of BiP binding (Fig. 1.11C) corroborates the results obtained in the immunoassays. Thus, Cdel construction seems to be less more affected than HC and Cmut. However it has to be noticed that values of BiP/HA ratio obtained in this case (Fig. 1.11C) were far lower than those obtained for NG mutated version (Fig. 1.6C), indicating that the degree of BiP binding is moderate, but not decisive, as in the case of NG.

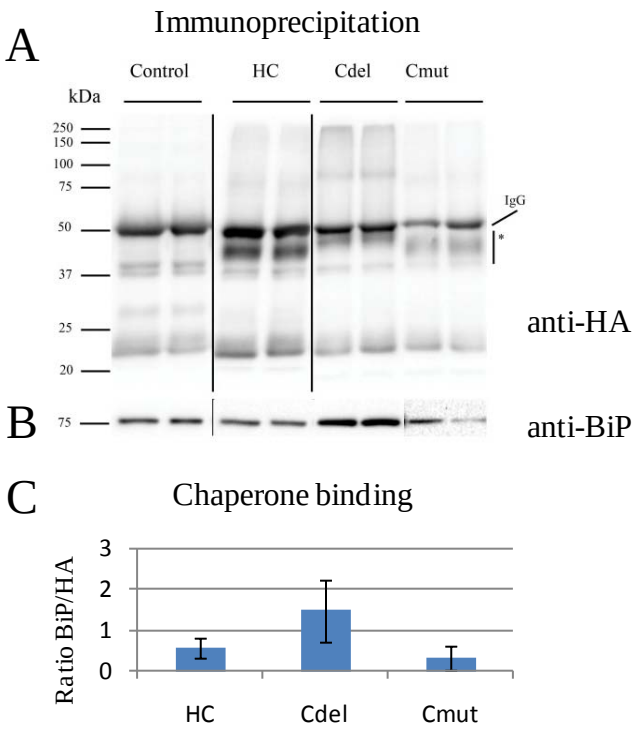

Figure 1.11: C terminus of HC isinvolved in proper folding of the protein. Protein extracts from total leaf extract from *Nicotiana* leaves expressing HC mutant versions were immunoprecipitated using HA agarose beads. As negative control, immunoprecipitated total extract from non transformed plant was used. Black line indicates different areas of same blotting experiment. A) Immunoblot with HA antibody. Asterisk (*) indicates immunoprecipitated protein. IgG, immunoglobuline G. B) Immunoblot with BiP antibody. C) Densitometric analysis of immunoblots A and B, to estimate the relative amount of BiP co- precipitating with Cdel and Cmut. The graph represents the ratio of BiP to HA-tagged protein (mean ± SE, n=24, p<0,001). Test ANOVA

1.5 Activity of HC requires N-glycosylation

Our experimental data indicated that N-glycosylation and an intramolecular disulphide bond formation are required for folding, and therefore ER export and trafficking of HC protein.

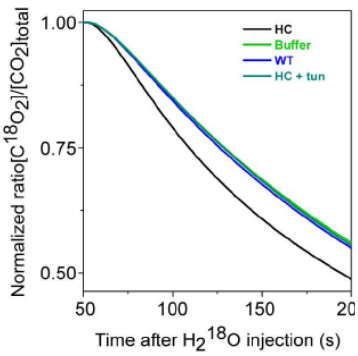

Figure 1.12: Activity of HC requires N-glycosylation. MIMS measurements of the change in $C^{18}O_2$ concentration as function of time after the 15 µL injection and mixing of air-saturated $H_2{}^{18}O$ (to final enrichment of 2,4%) into the 600 µL MS cell at pH 6,5 and 20ºC. Average traces of 3 repeats are presented, corresponding to HA immunoprecipitates from cells stably expressing HA tagged CAH1 (HC, black), HC treated with Tun (cyano), buffer (green), non transformed cells (WT, blue). From *Burén et al.,* 2011

In order to test the effects of these post-translational modifications on the activity of the CAH1 protein, in a collaborative work with a group of UPSC, it was applied a strategy that allowed the specifical measurement of the activity of CAH1, (see section 2.7.3 of Material and Methods). HC protein was immunoprecipitated from suspension culture cells stably expressing HC. The carbonic anhydrase activity of the immunoprecipitated HC protein was then measured as the decrease of CO_2 after injection of air-saturated labeled water ($H_2{}^{18}O$). The role of the N-linked glycans anchored to the HC protein was tested by performing the same experiment using immunoprecipitated HC protein from culture cells treated with tunicamycin for 24 h, a condition that mimics expression of NG isoform. The nonglycosylated HC isoform did not show any detectable activity (Fig.1.12 and Burén *et al.*, 2011) therefore clearly confirming, from another approach, that N-glycans are required not only for proper folding but also for activity of the protein. Nevertheless, these results should be confirmed *in planta.*

1.6 Complementation of knockouts of *Arabidopsis thaliana*

In order to confirm the observations, a new strategy using CAH1 knockout plants was performed. These plant lines show reduced starch accumulation and impaired photosynthesis phenotype when compared to the wild type (Burén et al., unpublished). To test whether the different CAH1 mutated versions would complement the starch and photosynthesis phenotype, *Arabidopsis thaliana* CAH1-knockout lines KO:361 and KO:029 (see section 2.3.3 of Material and Methods) were stably transformed by flowering dipping. For that, constructions previously described for transient transformation in *Nicotiana benthamiana* (see sections 1.2.3 and 1.4.1 in this Chapter) were used. That is, cultures of *Agrobacterium tumefaciens* strains containing NG, Cdel and Cmut mutated versions into the binary vector pmDC32. This experiment is still ongoing, but the results will help to understand the implications of the posttranslational modifications and C terminus in proper CAH1 functionality.

CHAPTER 2: FUNCTION OF COMPLEX N-GLYCANS AND LOCALIZATION OF CORE α(1,3)FUCOSYLTRANSFERASES

FUNCTION OF COMPLEX GLYCANS IN *Arabidopsis thaliana*

N-glycosylation of proteins is an important PTM, known to affect protein folding and stability (Strasser *et al.,* 2004; Shental-Bechor and Levy, 2008), as previously mentioned. Alteration of N-glycan processing steps taking place in the ER results in defective protein function and lethality (Koiwa *et al.,* 2003). However, the precise role of plant late N-glycosylation in Golgi system that generates plant complex N-glycans, is still discussed (see section 3.1 of Introduction).

Kang and coworkers (2008) suggested a potential protective role of complex N-glycosylation under salt stress conditions in *Arabidopsis*. They found that the complex glycan less 1(*cgl1*) mutant, which did not display any clear alteration of its phenotype under normal growth conditions (i.e., non stressed conditions), presented defects in cell wall-biosynthesis and mitotic arrest in the root apical meristem when it was subjected to salt stress (Kang *et al.*, 2008, von Schaewen *et al.,* 2008).

It is well known that salt stress is a common abiotic stress that plants have to deal with in nature; as it influences growth and development, and trigger a series of physiological processes in order to re-establish ion homeostasis and water status (Cubero *et al.,* 2009; Schoberer and Liebminger, 2013).The influence on plant tolerance to osmotic stress caused by the lack of complex N-glycosylation (and/or the subsequent excess of mannose residues), or by the absence of specific residues, as α(1,3) linked fucose is discussed in the present work. To try to clarify this question, experiments using *Arabidopsis* mutants defective in N-glycosylation were performed. These mutant lines are a knockout in the two known core α(1,3) fucosyltransferases (Fig. I3k), ΔFT, and a GnTI (Fig. I3e) enzyme knockout, the *cgl1* mutant previously mentioned. ΔFT mutant line was generated in Prof. Samuelsson laboratory and kindly provided for perform this study. Different features oth these mutants were analyzed: Root growth, cell wall resistance to degradation and expression levels of the cell wall modifying N-glycoenzyme endo β(1,4) glucanase (KOR1), which was known to be affected by N-glycan structure alteration (Kang *et al.,* 2008).

2.1 Defective glycosylation mutants are affected in root growth under salt stress conditions

In first term, the effect of salt stress on plant growth was tested in *Arabidopsis* Columbia 0 (wild type/wt), ΔFT and *cgl1* mutants, by growing them under increasing NaCl concentrations (0, 50, 75, 100 and 150 mM). Root length of seedlings grown for 17 days in MS medium plates with the different treatments was measured. No significant phenotypical differences were observed in

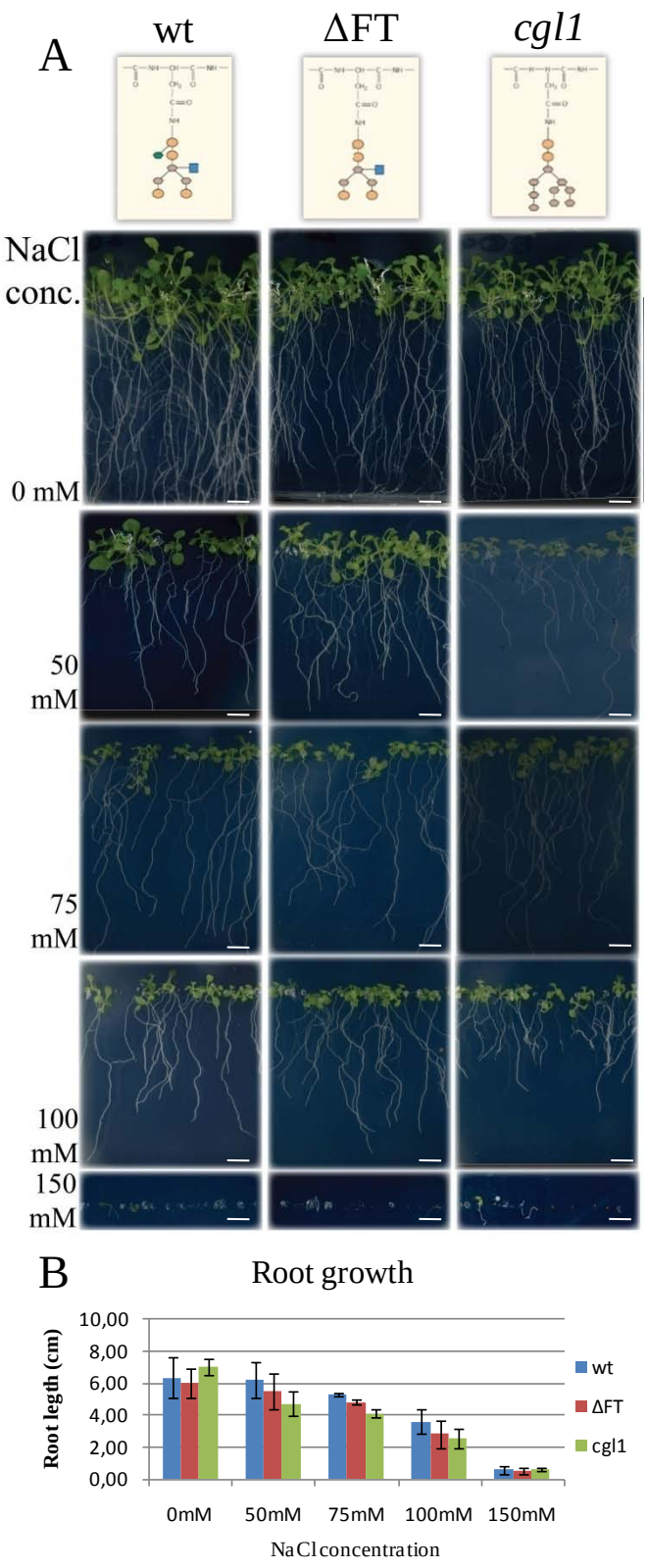

Figure 2.1: Plants with N-glycosylation defects show root growth arrest in salt stress conditions. A) Above, schematic representation of complex N-glycan structure in wild type (Col 0), ΔFT and *cgl1*. Green hexagon, α1,3 fucose, blue square, β1,2 xylose. Modified from Spremulli *et al.*, 2000. Below, 17 days old seedlings grown on MS-agar containing different NaCl concentrations. Scale bar 1 cm. B) Graph representing root growth versus NaCl concentration. Not statistical significance. (mean ± SD, n=60) Kruskal Wallis test.

control conditions (Fig. 2.1A), in agreement with previous studies performed by other research groups (von Schaewen *et al.*, 1993; Strasser *et al.*, 2004, Kang *et al.*, 2008).

A clear arrest of seedling development was observed with the highest NaCl concentration (150 mM) for the three lines, displaying no significant differences between them. However, slight differences were observed in intermediate NaCl concentrations (50, 75 and 100 mM). On the other hand, root growth of *cgl1* mutant seemed slightly more affected than that of ΔFT mutant, being the wild type the least affected. To confirm these observations, root length of each seedling was measured and a statistical analysis was performed, corroborating the differences observed by naked eye, that is, the effect on root growth was correlated with increasing NaCl concentration (Fig. 2.1B). Non statistical differences were appreciated when comparing root growth of wild type and mutant lines, however, a modest trend showing increased sensitivity to salt effect was inferred for the mutant lines with more marked N- glycosylation defects.

2.2 Assessment of cell wall consistency under the effect of degrading enzymes

Since defects in N-glycosylation are only noticed when plants grow under salt/osmotic stress conditions, revealed by the slightly higher sensitivity of mutants to salt treatments, it was decided to test whether cell wall was more susceptible to degradation in these defective N-glycosylation plant lines than in the wild type. As mentioned above and section 3.1.1 of Introduction, cell wall formation is closely related to the proper functioning of N-glycoslylation machinery, as some of N-glycosylated proteins, which specifically require those PTMs, are involved in its biosynthesis.

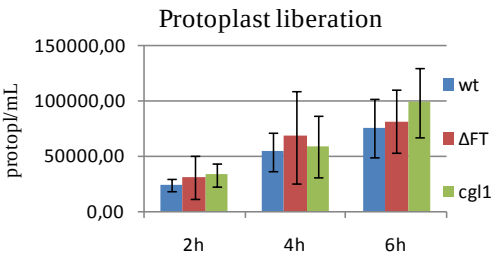

Figure 2.2: Amount of protoplast released at different times under enzyme degradation treatment. No statistical significance (mean ± SD, n=27). Test ANOVA.

Leaves of 3 weeks old wild type Col-0, ΔFT and *cgl1* plants were treated with a cell wall digestion solution containing cellulose and macerozyme (see section 2.3.1 of Material and Methods), and cell wall susceptibility was measured as number of released protoplasts at different time points. Number of protoplasts was estimated by counting them in an optical microscope (section 2.8.1 of Material and Methods). Consistent with the expectations, the number of released protoplasts was proportional to the exposure time of the leaves to the digestion solution, as higher number of protoplasts was released with increasing incubation lapses (Fig. 2.2). Although not statistically significant, a general trend, suggesting that mutants with impaired N-glycosylation where more sensitive to the digestion treatments than wild type was observed, since protoplast release was generally higher than that of wild type, especially when compared to *cg1l*. However, high variability of the absolute data hampered the attainment of clear significance in the statistical analysis.

2.3 Assessment of cell wall formation in cultured protoplasts

Due to lack of statistical significance of previous experiments, it was decided to try to detect possible cell wall alterations in N-glycosylation deficient plants from another approach. For that, it was tested whether the alterations in N-glycosylation could affect the timing of cell wall formation. Protoplasts of *Arabidopsis* wild type Col-0, ΔFT and *cgl1* lines were isolated from leaf tissue and subsequently incubated with different NaCl concentrations (0, 50, 75, 100 and 150 mM). Then, they were stained with the cell wall labeling agent Calcofluor white. Calcofluor fluorescence was measured in a plate reader with an appropriate filter for different time lapses, in order to monitor possible differences in the rate of cell wall formation among the cell lines and/or

salt treatments. Unfortunately, no fluorescence emission of Calcofluor different of the background signal, could be detected in any cell line or salt treatment (data not shown).

Therefore, a different approach was implemented. In this occasion, *Arabidopsis* protoplasts of Col-0, ΔFT and *cgl1* were isolated from leaves tissue, washed and cultured. Subsequently, they were stained with Calcofluor, in order to be visualized at different time lapses with a fluorescence microscope. Protoplast culture is challenging, and different medium might affect cell wall formation timing. Therefore, in order to find the most suitable for cell wall formation assessment, the assay was performed in three different media: W5, W5a and a modified-*Medium B* used in Kwon *et al.* 2005 (see section 2.3.1 of Material and Methods). Cultures were maintained for 48h. Although protoplasts seemed healthy, Calcofluor fluorescence could not be appreciated even after 48h, neither differences between wild type and mutants (Fig. 2.3). This absence of indicative Calcofluor fluorescence was consistent with previous data from plate reader, which were obtained in even shorter time periods. Similar experiments by other research groups described a similar lack of calcofluor fluorescence after seven days of protoplast culture (Abreu *et al.*, personal communication).

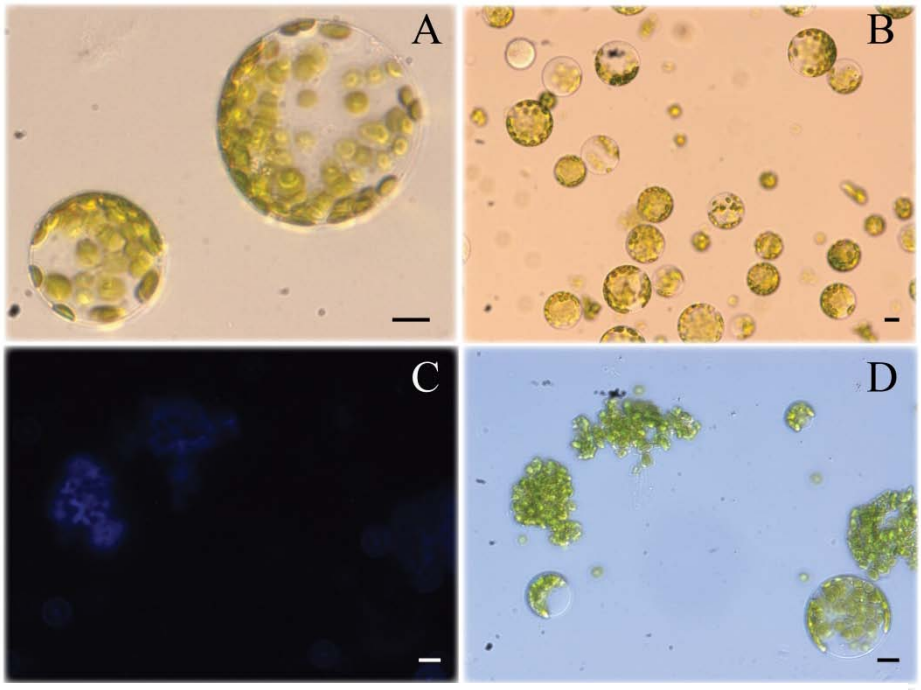

Figure 2.3: Released protoplasts observed at microscope. Different views of protoplast of *Arabidopsis* wild type line in W5 buffer stained with Calcofluor, at bright field filter (A, B and D), or DAPI filter (C) Scale bar 5 µm in A, C and D; 10 µm in B

2.4 Endo β (1-4) glucanase expression levels under salt stress conditions are highly variable

Cell wall formation is controlled by different enzymatic activities, some of them carried out by N-glycosylated proteins, as KOR1 (AT5G49720). It has been reported that PTMs, as N-

glycosylation, influence in cellulose biosynthesis and cell wall functionality under salt stress conditions (Kang *et al.,* 2008, and section 3.1 of Introduction).

In order to discriminate whether ΔFT and *cgl1* mutants presented differences in osmotic stress tolerance due to defects in KOR1 entire complex N-glycosylation, or exclusively in fucosylation, it was attempted a strategy that would help to understand the extent of N-glycosylation influence on KOR1 activity and gene expression. Additionally, it would clarify whether the differences in osmotic stress tolerance are due to general defects in Golgi complex N-glycosylation or in fucosylation exclusively. It was hypothesized that expression levels of the corresponding *kor1* gene might be induced, due to depleted protein activity in N-glycosylation defective mutants and/or after salt treatment. Thus, KOR1 expression was compared in *Arabidopsis* wild type Col-0, ΔFT and *cgl1* mutants exposed to different salt concentrations (0, 50, 75, 100 and 150 mM). RNA extraction of seedlings grown in MS plates with the different salt treatments was followed by RT-PCR expression analysis (see section 2.2.3 and section 2.2.9 of Material and Methods). RNA isolation from plants corresponding to the 150 mM NaCl treatment was not possible due to the extreme seedling alterations and RNA degradation (not shown). As a RT-PCR housekeeping control constitutively expressed ADP ribosylation factor gene (ADP-RF, AT3G62290) was used (Craciun *et al.,* 2006). After 35 amplification cycles, amplification of both *kor1* and *adp-rf* genes was successful, as shown by corresponding bands at the expected size (800 bp and 400 bp, respectively) in an agarose gel (Fig. 2.4B). After image analysis of 4 different gel replicates, pixel density was normalized to intensity of housekeeping gene, in order to accurately assess the expression level of *kor1* as described in section 2.9 of Material and Methods. The high variability of the data impeded

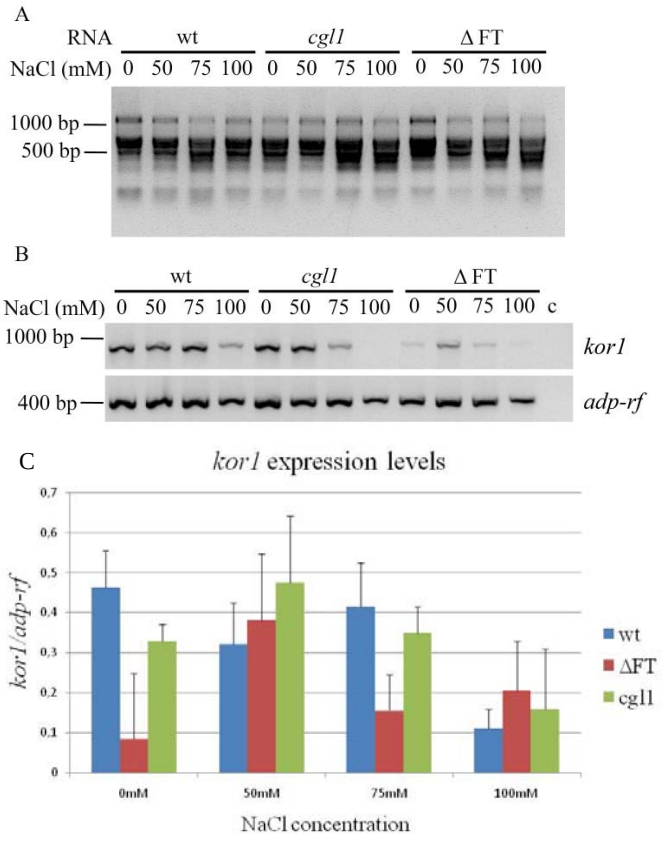

Figure 2.4: Endo β (1,4) glucanase expression levels show high variability regardless the impaired N-glycosylation and salt treatments. A) RNA control of samples from *Arabidopsis* lines, under increasing NaCl concentration growth conditions. Pattern show good quality of samples. 900 µg of RNA per line. B) RT-PCR of *kor1* gene. Constitutive *adp-rf* gene was used as control. Reaction without RNA sample was used as negative control. C) Densitometric analysis shows high variability in expression levels of *kor1*, and there are no statistical significant differences between wt and mutants and salt treatments. (Mean ± SE, n=48) Test Kruskal Wallis. Wt, wild type.

a statistically significant conclusion, neither was possible to estimate any relation between expression levels and salt concentration in the media (Fig. 2.4C). In any case, we did not found significant differences in *kor1* expression that might be related to the absence of α(1,3)fucose (ΔFT mutant) or the whole N-complex glycan (*cgl1*) suggesting a different mechanism to compensate the potential protein activity decrease in these mutants, as for example posttranslational activation.

LOCALIZATION OF CORE α(1,3) FUCOSYLTRANSFERASES

Plant α(1,3)fucosyltransferases (FTs) activity consists in the addition of a fucose moiety through α(1,3) linkage to the proximal N –linked N-acetylglucosamine (GlcNAc) of the standard ER glycan (Figs. I3k and I4). As described in the previous section, the presence of α(1,3)fucose does not seem to be significantly relevant for proteins targeted to the cell wall. However, this is the only step in the N-glycosylation pathway catalyzed by two enzymes acting in the N-gycan core: α(1,3)fucosyltransferase 11 (FT11) and α(1,3)fucosyltransferase 12 (FT12). As mentioned in section 4.4 of Introduction, it is still controversial whether both isoforms function redundantly or there is a differential localization pattern in diverse Golgi stacks, leading to fucosylation of specific sets of subcellular proteins with differential targeting.

Previous confocal microscopy observations by Forth et al. (unpublished), showed Golgi localization of FTs, as expected from their post-ER N-glycan modification function. In addition, a partially redundant role of both FT isoforms similar to that described by other authors (Strasser et al., 2004) was found. Indeed, FT11 knocking out did not seem to affect the fucosylation pattern of any of the sub-cellular set of proteins analyzed. However, FT12 activity appear to play a predominant role in the fucosylation of plastid N-glycoproteins, with little effect on the fucosylation of other sub-cellular fractions (Forth et al., unpublished).Taking into account the finding of CAH1 in the plant chloroplast, it was hypothesized that there might be different specialized Golgi compartments (dictiosomes or stacks) carrying different glycosyltransferases and that FT12 might be involved in the fucosylation of glycoproteins trafficking through the endomembrane system to the chloroplast.

Hence, as part of the collaboration with Prof. Samuelsson´s group (UPSC, Sweden), a different approach was attempted with the aim to elucidate the exact subcellular localization of FT11 and FT12. This would help to determine whether they colocalize or exist in different Golgi stacks, which would indicate the existence of differential protein processing.

2.5 Selection of lines with high expression levels of myc and HA tagged fucosyltransferases

Arabidopsis lines tagged with *myc* and *hemmaglutinine (HA)* epitopes in FT11 and FT12 respectively, in a wild type background, were generated in Prof. Samuelsson laboratory and kindly provided in order to monitor FTs subcellular localization by immunohistochemistry assays. To assess the localization of both FTs within the same plant, seedlings expressing both tags were selected. As the objective of the present work was the accurate localization of the FTs within the Golgi stacks by immuno-detection assays, it was also crucial to select plants that expressed both tagged proteins properly. A screening with *Arabidopsis* plants from different 2x7 and 1x6 lines expressing FT11-myc and FT12-HA proteins was performed. Protein extracts of different rosette leaves were analyzed by western blot, and plants with the highest expression levels of the tagged FTs in total leave tissue were selected (2x7:9 and 1x6:8, Fig.2.5).

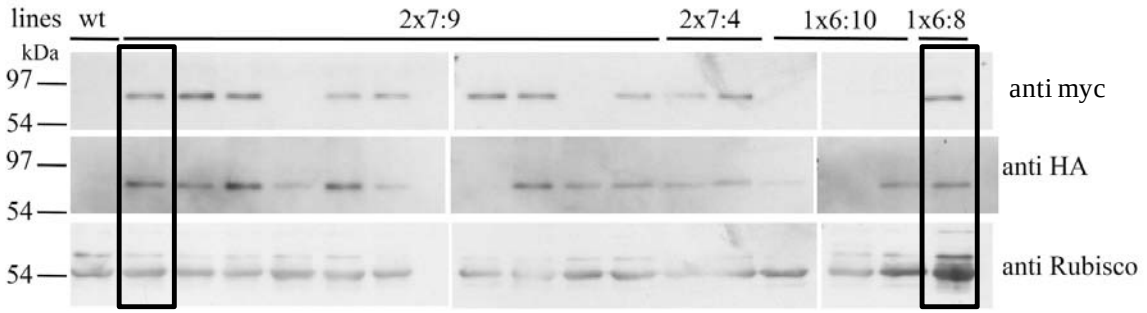

Figure 2.5: Screening of *Arabidopsis* seedlings containing tagged fucosyltransferases. Total leave extract of *Arabidopsis* wild type and mutant lines containing tagged versions FT11-myc and FT12-HA. Selected plants are marked with a black rectangle. 10 µg of protein per line

2.6 Expression levels of tagged proteins in seedligns of selected *Arabidopsis* lines are higher in early stages of development

In order to determine in which developmental stage of the seedlings the expression levels of the tagged proteins was highest, *Arabidopsis* selected homozygous lines (Fig. 2.5), expressing both FT11-myc and FT12-HA were grown for 1 to 4 weeks in standard conditions, as described in section 1.2 of Material and Methods. Protein extracts of rosette leaves were

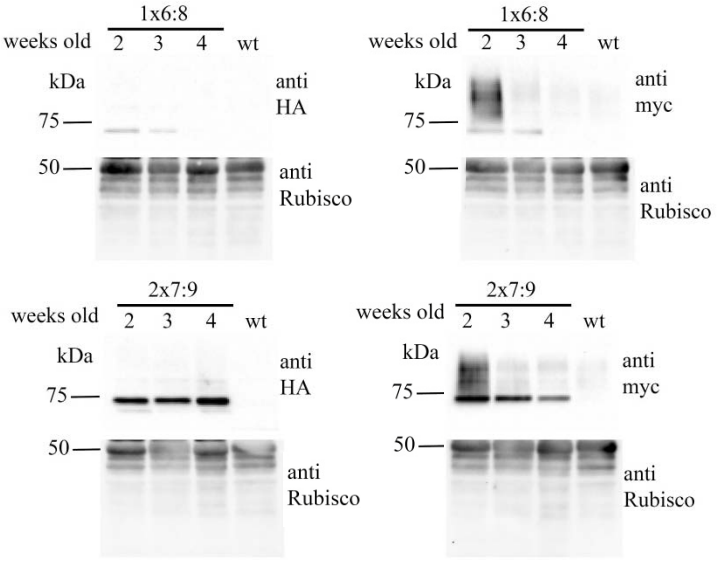

Figure 2.6: Expression levels of FT11-myc and FT12-HA are higher in early stages of development of seedlings. Total leaf extract from *Arabidopsis* lines analyzed with epitope-tag and Rubisco antibodies- 10 µg of protein from a pool of 5-6 plantules a per line.

obtained, and western blot analysis was performed with anti-HA, anti-myc and anti-Rubisco antibodies, the former one used as a loading control (Fig. 2.6). Immunodetection with anti-myc antibody showed decreasing expression levels with time, being two weeks the stage with highest amounts of recombinant protein. This effect was also observed in the analysis with anti-HA antibody, but to a lesser extent. Therefore, two weeks after sown was selected as the best moment for further microscopy immunolocalization experiments.

2.7 Golgi localization of FT11 and FT12

As previously mentioned, the study performed by Forth et al (unpublished) reported FTs' localization in the Golgi apparatus. Nevertheless, in order to discern whether FT11 and FT12 co-localize in the same Golgi stacks or they are separated, it was decided to endeavor a more specific localization analysis with Transmission Electron Microscopy (TEM). As a complementary and guiding technique, an approach based in cell fractionation and further FT localization, based in colocalization with different organelles markers, was carried out. For that, microsomes of selected lines (one of 2x7:9 and other from 1x6:8, see Fig. 2.5) were isolated and further separated in sucrose gradients as described in section 2.1.2 and 2.1.3 of Material and Methods. Microsomes are the membranous fraction obtained from total leave extracts (TE), widely used for study of membrane bound proteins. TE is fractionated, separating most of the endomembrane system vesicles and cisternae, including ER, Golgi, vacuole, lisosomes, and the rest of organelles, from a soluble fraction containing mainly the hydrophillic cytosol and organelles' aqueous media. Both tagged proteins were mainly found in the microsome fraction (M) (Fig. 2.7), when compared to the total extract (SN1) and the soluble fraction (SN2), indicating that the microsome isolation was performed properly.

After isolation and concentration of microsomal fraction, it was loaded onto sucrose gradients. These gradients allow separation of different endomembrane vesicles according to density (Raikhel and Chrispeels, 2000). Unfortunately, a specific anti-Golgi antibody was not available, therefore hampering the identification of those specific fractions. Instead, other antibodies labeling other fractions helped to approximate the localization of tagged proteins. Unfortunately, no clear signal indicative of distribution of the myc and HA tagged FT along the sucrose gradient was obtained (data not

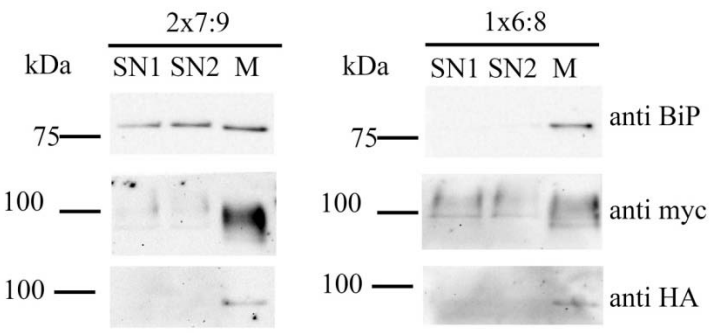

Figure 2.7: HA and myc tagged FTs are mainly found in microsome fraction. Total extract, soluble and microsome fractions of lines 2x7:9 and 1x6:8 are shown. 5 µg of. protein per line for BiP control; 20 µg of protein for HA and myc blots. SN1 and SN2, supernatants 1 and 2, respectively; M, microsome fraction.

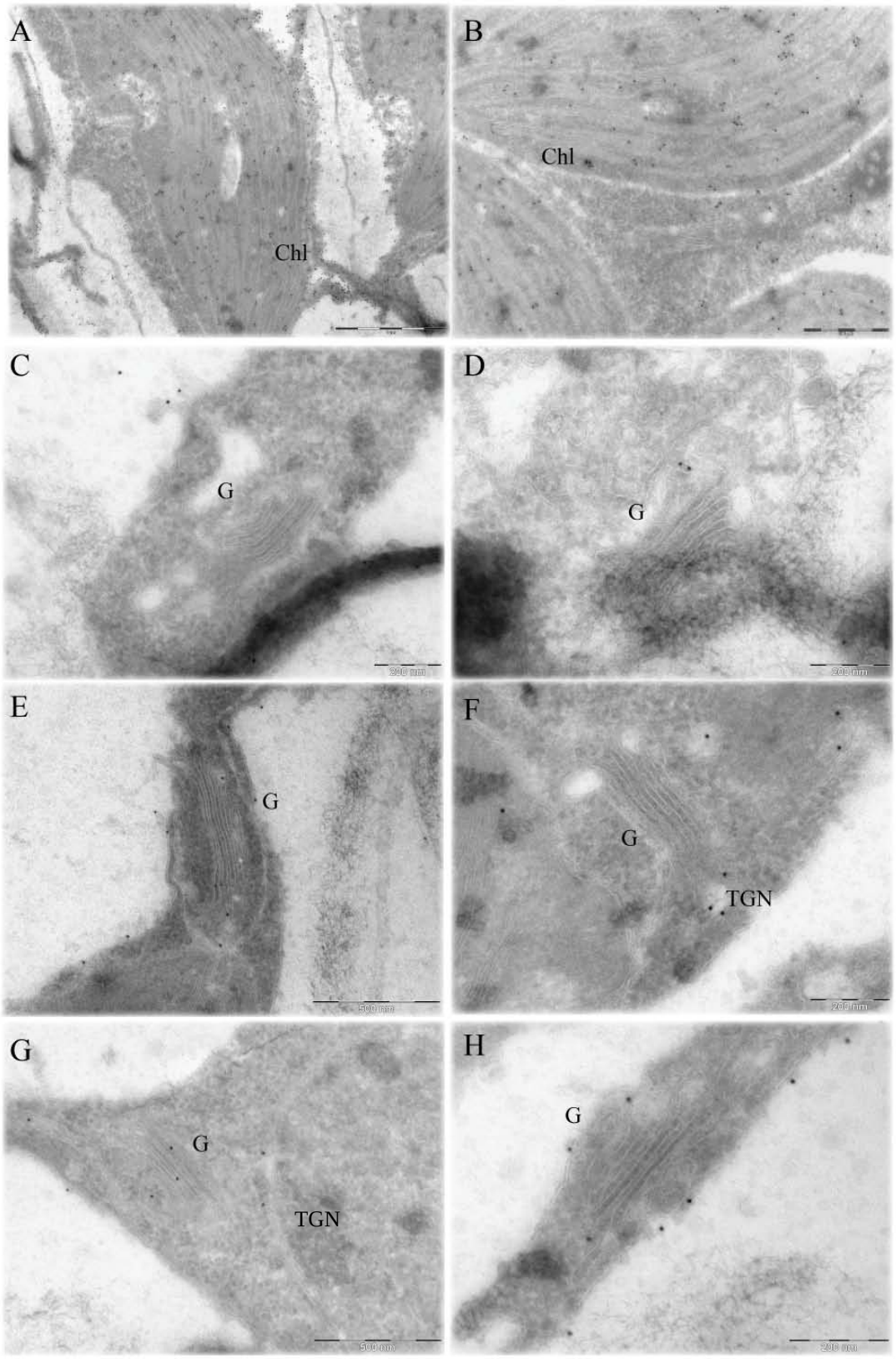

Figure 2.8: Transmission Electronic Microscopy from cryosections of *Arabidopsis thaliana* mesophyll cells A-B) General vision of the cell. Line 2x7:9 with anti myc 1:10. C-H) Zoom focusing in individual Golgi apparatus. C) Line 1x6:8 with anti myc 1:500. D) Line 1x6:8 with anti myc 1:1000. E and F) Line 1x6:8 with anti myc 1:500 G) Line 1x6:8 with anti HA 1:25. H) Line 1x6:8 with anti myc 1:500. Black dots indicate gold particles of tagged FTs with HA or myc. G, Golgi apparatus; Chl, chloroplast, TGN, trans-Golgi Network

shown). Therefore, this technique was not accurate enough for determine FT 11 and 12 localization, neither discern whether they were placed in the same or different Golgi stacks, making necessary to perform a more specific subcellular localization analysis.

Therefore, immunolocalization of tagged proteins by TEM was performed. Two and four weeks old *Arabidopsis* seedlings from wild type Columbia-0, and 2x7:9 and 1x6:8 mutant lines were tested, as accuracy of the technique maybe allow to notice any difference of FTs location according to plant's age. Leaf material soaked in a parafolmaldehide fixative solution, cryofixed and cut in ultrathin sections. The sections were laid on special TEM grids and subjected to immunocitochemistry, in order to incubate with primary and secondary antibodies linked to gold particles, for proper TEM visualization (see section 2.8.4 of Material and Methods). Different dilutions of primary antibody were tested, according to empiric experience and manufacturer's instructions, in order to obtain the best resolution and signaling for both myc and HA tagged proteins. Such dilutions were 1:100 and 1:25 for HA antibody, and 1:1000, 1:500 and 1:200 for myc antibody. The objective was, in first term, to set up the best conditions for primary and secondary antibody dilutions for each epitope, and subsequently to label simultaneously FT11-myc and FT12-HA with secondary antibodies with gold particles of different sizes, in order to allow direct *in cell* comparison, and then distighuish the effectively localization of each FT. However, in those individual preliminary tests, in all cases, and in both mutant lines, a homogeneous distribution of spots corresponding to gold particles was observed, in areas corresponding to cytoplasm, chloroplast membranes, occasionally some Golgi stacks and in vesicles of the trans Golgi network (Fig. 2.8). All tested dilutions presented a moderate and disperse labeling that was therefore considered as background. General low and in some cases unspecific levels of gold particle labeling were observed. The absence of a differential labeling, or in any case, a determinate localization pattern of gold particles, made it challenging to conclude or hypothesize any conclusion about individual localization of FTs. Then, it was no feasible to perform a double immunolocalization due to the high risk of non significant results. It is probable that these proteins are not highly represented in the cell, hampering to unequivocally locate them in small structures as the Golgi stacks.Additionally, Golgi structures are quite difficult to find and discriminate in ultrathin cutting as reported in several studies, making FTs subcellular localization within Golgi cisternae quite challenging (Han *et al*., 2013).

CHAPTER 3: IDENTIFICATION OF NEW N-GLYCOSYLATED PROTEINS TARGETED TO CHLOROPLAST THROUGH THE ENDOMEMBRANE SYSTEM

Several proteomic studies had shown the occurrence of proteins in the chloroplast of higher plants without transit peptide (TP) (Kleffmann *et al.,* 2004; Friso *et al.,* 2004, Rutschow *et al.,* 2008; Zybailov *et al.,* 2008; Armbruster *et al.,* 2009), suggesting that the targeting process is more complex than previously thought (section 1 of Introduction). One of the first evidence of an O-glycosylated protein in *Pisum sativum* (pea) chloroplasts was reported by Gaikwad *et al.,* in 1999. After the discovery of a glycoprotein trafficking pathway from the endomembrane system to the chloroplast, for proteins bearing complex N-glycan, by Villarejo and co-workers (2005), occurrence of some N–glycoproteins were reported in monocot chloroplasts: α-amilase I-1 and nucleotide pyrophosphatase/phosphodiesterase 1(NPP1) were found in *Oriza sativa* (rice) chloroplasts (Asatsuma *et al.,* 2005; Nanjo *et al.,* 2006). Experimental data showed that NPP1 most probably harbors only high mannose *N*-glycans (Nanjo *et al.,* 2006).
An attempt to localize NPP1 in *Arabidopsis* chloroplasts was performed in a collaborative work with Prof. Samuelsson's group. However, it was no possible to find the protein in its final location, thus it had to be set a different strategy, in order to find a model protein which allow us to continue investigating this issue.

The presence in the chloroplast of N-glycoproteins harboring only high-mannose-type N-glycans (HMGP) acquired in the ER, like rice NPP1, raise the possibility that those proteins could be transported directly from the ER to the chloroplast. Indeed, direct transport of molecules between these two organelles is already demonstrated, but regarding lipid precursors (Wang and Benning, 2012; section 4.1 of Introduction). Although such lipid precursor's traffic has been characterized, some studies speculate with the possibility of a direct transport of proteins between ER and chloroplasts (Andersson *et al.,* 2007; Bhattacharya *et al.,* 2007; Tan *et al.,* 2011; Krause *et al.,* 2012; Gagat *et al.,* 2013).Therefore, it is possible that HMGP could reach the chloroplast using or sharing components of the pathway for lipid precursor's transport.
In *Arabidopsis thaliana*, a series of mutants, named *tgd* (tri-galactosyl diaciylglycerol), have been characterized. They are defective in the expression of different key proteins involved in the lipid transfer process between the ER and the chloroplast (Fig. I5). *tgd1, tgd2* and *tgd3* are defective in the expression of proteins from the inner envelope membrane translocon complex, and *tgd4* is defective in the expression of an ER resident protein, that seems to mediate lipid transfer between this organelle and the outer plastid envelope membrane (Xu *et al.,* 2008).Thus, the mutants have disrupted the import of lipid precursors from ER into the chloroplast.

Then, if some proteins are transported sharing some components of this pathway, the *tgd* mutants constitute a proper tool for ER-chloroplast protein traffic analysis.

3.1 The *Arabidopsis tgd1* and *tgd4* mutants are phenotypically different than the wild type

*tgd1*and *4* mutants were kindly provided by Prof. Christoph Benning from Michigan State University (USA). *tgd1* is defective in the expression of a permease-like protein, which belongs to TGD traslocon, and *tgd4* is defective in the expression of an extraplastidial protein located in ER (Fig. I5 and section 4.1 of Introduction). These mutants are representative of the two steps in the lipid precursors transport pathway. Phenotypic differences were observed during plant development in soil grown plants, as the two mutants exhibited a slower growth rate and slightly paler green colour than the wild type (wt) (Fig. 3.1). Curiously, *tgd4* mutant presented the more differential phenotype (Fig. 3.1 C and D).

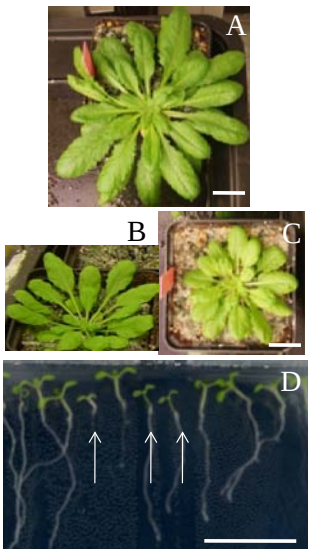

Figure 3.1: TGD mutants present a distinctive phenotype. Differential phenotype observed in 3 weeks old mutants *tgd1* (B) and *tgd4* (C) when compared to wild type (A), presenting slower growth and paler color in plants grown in soil at standard conditions. D) Screening for selection of homozygous *tgd4* mutant lines. Seedlings grown in MS media for 1 week. Arrows show homozygous seedlings. Scale bar 2 cm.

Due to infertility of homozygous lines of *tgd4*, lines were maintained in heterozygosity (Xu *et al.*, 2008), and homozygous specimens had to be selected from screening the offspring of heterozygous plants. For that, seeds were germinated in agar plates, where phenotypic differences were noticed at naked eye (Fig. 3.1D) and further confirmed by PCR (Fig. 3.2), as described in detail in section 2.2.8 of Material and Methods. Briefly, DNA of each plant was subjected to two PCRs, one with primers amplifying specifically the wild type TGD4 gene, and the other amplifying a fragment of the T-DNA insert placed in the interrupted TGD4 gene (Fig. 3.2A). Homozygous lines were those presenting a 750 bp band corresponding to the T-DNA insert, and no band in the PCR amplifying the 1000 bp wild type gene (Fig. 3.2B). After selection of homozygous plants, protein extracts were isolated and stored frozen for further analysis.

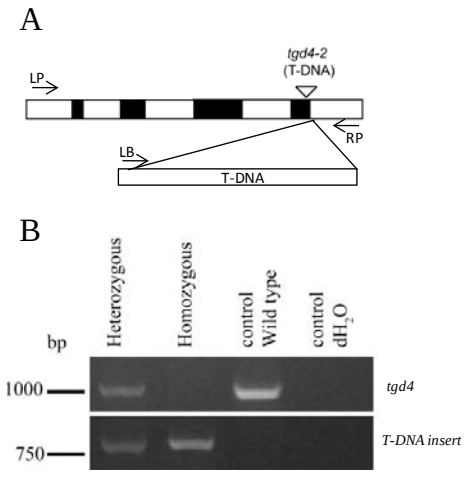

Figure 3.2: Screening of TGD4-2 mutant. A) Schematic representation of location of T-DNA insert in the *tgd4* gene sequence. Introns are marked as black boxes. Primers used in each screening reactions are marked. Adapted from Xu *et al,* 2008. B) PCR reactions amplifying wild type *tgd4* gene (upper panel) and a fragment of the gene containing the T-DNA insertion (lower panel).

3.2 *tgd* mutants are not significantly different than wild type at protein level, and routes for protein targeting to the chloroplasts are independent of the lipid precursor transport pathway

In order to rule out a possible alteration of general cellular processes that could eventually preclude the use of the *tgd* mutants for the planned proteomic studies, a search for general variations in protein levels was carried out. This study was followed by more specific analyses to test whether *tgd* mutants were affected in protein transport pathways to the chloroplast.

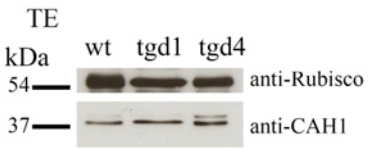

Figure 3.3: There are not significant differences in *tgd* mutants and wild type protein levels at the cellular level. 10 µg of protein per line. TE, total extract.

Proteins from leaf total extract of wild type, *tgd1* and *tgd4* mutants were analyzed by SDS-PAGE and western blot, using anti-Rubisco and anti-CAH1 antibodies, as representative markers for the canonical and endomembrane chloroplast trafficking pathways respectively (Fig. 3.3). No significant differences were observed between mutants and wild type samples in Rubisco or CAH1 immunodetection, indicating that there were no general metabolic failures affecting the accumulation of these proteins in the *tgd* mutants. Therefore, the *tgd* mutations do not seem to hamper the stability of selected protein markers, and could be used for further analyses.

In order to study whether the *tgd* mutations affect the canonical or endomembrane protein pathways specifically to the chloroplast, cellular total extract and purified chloroplasts (Chl) from wild type, *tgd1* and *tgd4* mutants were subjected to western blot analysis, using the same antibodies (anti-Rubisco and anti-CAH1) as markers. Nevertheless, in first term, the purity of chloroplasts fractions was checked. For that, aliquots were taken in each step of the chloroplast isolation process, demonstrating that main contaminants of other organelles remained in discarded fractions and concomitant enrichment of chloroplast proteins was achieved (Fig. 3.4). It can be appreciated that total chloroplast preparations present a low degree of contamination, as slight levels of ER and vacuole markers can be detected in he obtained total chloroplast fraction, although a general decrease has taken place in comparison with initial fractions. The occurrence of a low degree of contamination is common and widely reported in

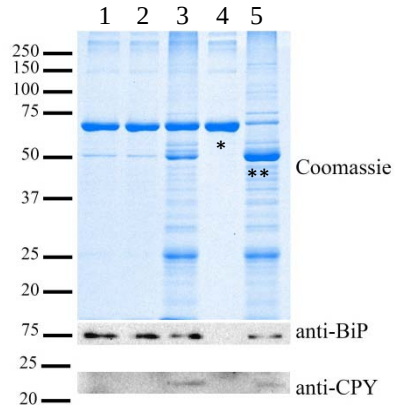

Figure 3.4: Isolation of total chloroplast fractions: Samples taken in different steps of the chloroplast purification process were analyzed by Coomassie blue staining and immunoblotusing anti-BiP and anti-CPY antibodies. 15 µL of each fraction per lane. *BSA; ** Rubisco; arrow indicates vacuolar processed form of CPY. 1, Homogenized leaves total extract; 2, soluble cellular fraction; 3, crude chloroplast and other membranous cellular components; 4, wash of intact chloroplasts, containing mainly broken chloroplasts; 5, intact total chloroplasts fraction.

biochemical fractionation experiments (Huber *et al.*, 2003).

In order to test the specificity of the antibodies/lectin against complex and high mannose type N-glycans, a prior test was performed at total extract level on *Arabidopsis* samples. Wild type (Col-0), *Complex glycan less 1* mutant, *cgl1*, and one of the *tgd* mutants (*tgd4*) were compared by western blot. *cgl1*was used as control, due to its inability for further modification of N-glycans in Golgi system, and the consequent accumulation of proteins harboring only high mannose type glycans. As expected, the wild type and *tgd4* mutant showed bands corresponding to glycoproteins of different molecular mass with the anti-β(1,2) xylose antibodies (Fig. 3.5C). In contrast, there was no labeling of *cgl1* protein extracts, indicating a complete absence of complex N-glycans in the proteins of this mutant. Concomitantly, a higher level of high mannose type glycoproteins, displayed by a stronger Con A labeling, was found in *cgl1*, unlike wild type and *tgd4* mutants samples, which did not show such accumulation (Fig. 3.5B).

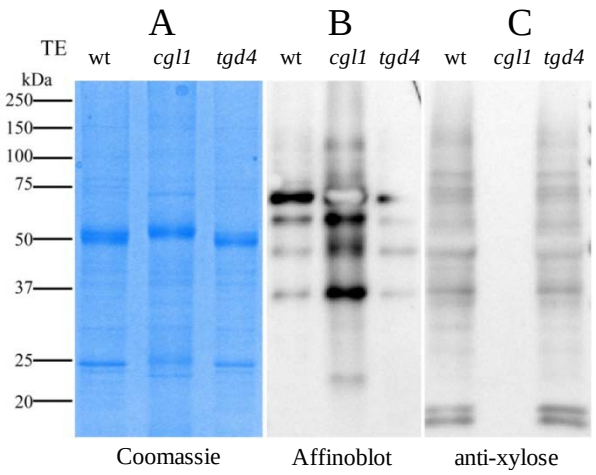

Figure 3.5: Control of specificty of ConA lectin and anti-xylose antibody. Total leave extracts of *Arabidopsis* subjected to affinoblot (B) and westen blot (C). Coomasie blue staining (A) was used as a loading control. 5 µg of protein per lane. TE, total extract

Once purity of samples and the immune specificity were tested, analysis of total extracts and chloroplast fractions of wild type, *tgd1* and *tgd4* were performed. Enrichment in Rubisco (Fig. 3.6B) and CAH1 (Fig. 3.6C) in all chloroplast preparations compared to the total extract was observed, both in wild type and mutant samples, checking, in first term, that isolation process of

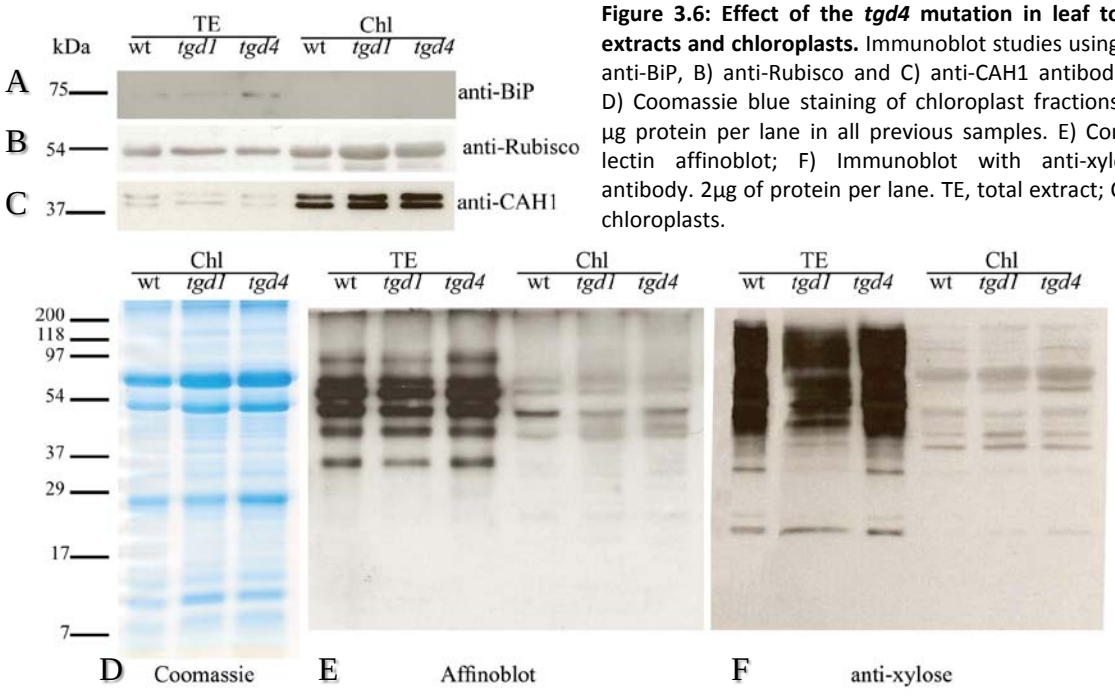

Figure 3.6: Effect of the *tgd4* mutation in leaf total extracts and chloroplasts. Immunoblot studies using A) anti-BiP, B) anti-Rubisco and C) anti-CAH1 antibodies. D) Coomassie blue staining of chloroplast fractions. 5 µg protein per lane in all previous samples. E) Con A lectin affinoblot; F) Immunoblot with anti-xylose antibody. 2µg of protein per lane. TE, total extract; Chl, chloroplasts.

the organelles has been successful. There were no differences in protein pattern when comparing wild type and mutant samples, indicating that the machineries involved in the targeting and import of these proteins to the chloroplast were not affected by the TGD1 and TGD4 specific mutations. The level of ER contamination of the chloroplast fractions was significantly reduced, as shown by the moderate signal of anti-BiP antibodies in all samples (Fig. 3.6A). The presence of N-glycoproteins with complex-type N-glycans in *Arabidopsis* chloroplasts was demonstrated by immunoblot using anti-β(1,2) xylose antibodies (Fig 3.6F) confirming previous reports (Villarejo *et al.,* 2005). Aditionally, no significant differences in the amount of these specific N-glycoproteins among wild type, *tgd1* and *tgd4* chloroplasts fractions was found. This result, together with that of CAH1, indicated that the pathway to import N-glycoproteins bearing complex N-glycans from endomembrane to chloroplast is not affected by the *tgd* mutations. Affinoblot analysis with Con A lectin showed indeed the presence of high mannose N-glycoproteins in chloroplasts from all three samples (Fig. 3.6E). It is not surprising to find high mannose type glycoproteins in the wild type and *tgd4* chloroplast fractions, since proteins of this kind have been previously found in *Oriza sativa* plastids, although the trafficking pathway remains unknown (Nanjo et al., 2006). The protein pattern observed in the affinoblot with Con A lectin was different to the immunoblot with anti-β(1,2) xylose, indicating that proteins labeled by both methods were probably different. However, no significant changes in the level of plastid high mannose type N-glycoproteins was observed between wild type and *tgd* mutants, probably due to the complex protein pattern.

3.3 The transport of high mannose type N-glycoproteins to the chloroplasts is affected in the *tgd4* mutant but not in *tgd1*.

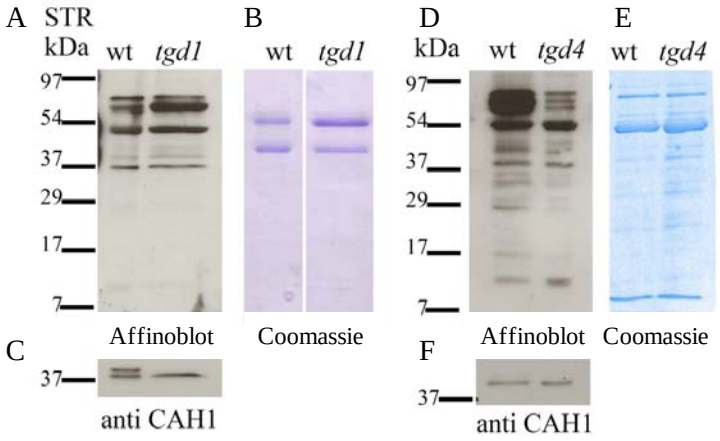

Figure 3.7: High mannose type N-glycoproteins levels are not altered in *tgd 1* but affected in the *tgd4* mutant, in stroma samples. Protein patterns of purified STRs are shown. A and D) Affinoblot. 10 µg of protein per line, B and E) Coomassie, 5 ug prot per line, C and F) CAH1, 10 µg of protein per line. STR, stroma

As no significant differences in N-glycoproteins levels were observed in total chloroplast fractions, stroma purification was performed as described in section 2.1.1 of Material and Methods. Stroma preparations exhibit less complex poplypeptide patterns, as they are free of membranes, thus offering a simpler model for protein

pattern studies. Immunoblot analysis of possible contaminants showed that stroma purification decreased the BiP signal compared to total chloroplast fraction for all the samples studied (Suppl. Fig. S3.1, upper panel), meaning that most of the ER proteins were removed in the stromal fractions. Rubisco level, however, remained constant as expected for a soluble stromal protein, indicating that the stroma purification was indeed successful (Suppl. Fig. S3.1, lower panel).

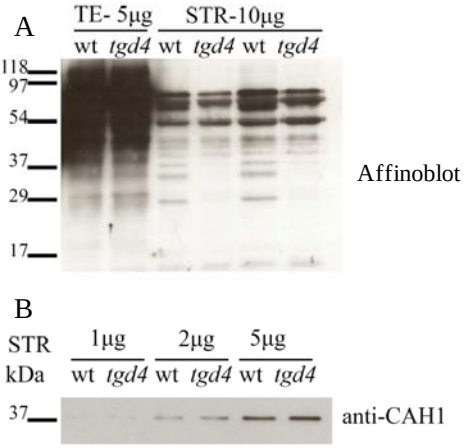

Figure 3.8: Differential high mannose N-glycans proteins (HMGP) pattern of *tgd4* mutant. Increasing amounts of stroma samples of wt and *tgd4* subjected to affinoblot (A) and immunoblot against CAH1, used as control (B). STR, stroma, TE, total leaf extract

In purified stroma samples, Con A lectin signal was slightly variable in *tgd1* sample when compared to wild type (Fig. 3.7A). However, more striking differences were observed in the *tgd4* high mannose protein pattern (Fig. 3.7D). This fact, together with the noticeable differential phenotype of *tgd4* mutant (Fig. 3.1C and D), and the putative role of TGD4 protein in the setup of plastid associated membranes (PLAMs), were decisive to select this mutant for perform proteomic analysis. To assess if there was any trend regarding high mannose glycoproteins in the different mutants, several different isolations were carried out. The observation consistently repeated was that a lower level of high mannose glycosylation was exhibited in the *tgd4* mutant proteins when compared to the wild type extracts (Fig 3.8). This fact points out to an alteration of the high mannose N-glycoproteins accumulation in the stroma of *tgd4* mutant. Taken together, differential HMGP pattern support *tgd4* as the more suitable mutant for perform proteomic analysis. These results suggest that the targeting of stromal high mannose glycoproteins is somehow linked to the activity of a functional TGD4 protein and a defective protein prevents these proteins to reach the chloroplast in similar levels than wild type or *tgd1*.

3.4 Identification of N-glycosylated proteins by different methods

Once *tgd4* mutant was chosen as the best tool to identify new HMGP in *Arabidopsis* chloroplasts, different approaches comparing proteome composition with the wild type were performed.

3.4.1 Two dimension differential gel electrophoresis (2D-DIGE)

Initially, a comparison of *Arabidopsis* wild type and *tgd4* stromal protein content was performed using two dimension differential gel electrophoresis (2D-DIGE). This technique allows visualization differential protein expression in two different samples (see section 2.4.4 of Material and Methods for detailed protocol). Pure stroma fractions previously purified of wild type and *tgd4*, and a mixture of both, were labeled with different fluorescent dyes (Cy5 and Cy3, alternatively for *tgd1* and *tgd4* and Cy2 for the mixture). After mixing Cy5, Cy3 and Cy2 labeled samples, an isoelectro-focusing gel followed by a second dimension electrophoresis were performed. Fluorescence was detected in Proteomics Facility of the Scientific Park of Madrid, using a Typhoon 9400 (GE Healthcare), selecting 532/580, 633/670 and 488/520 nm wavelengths for Cy3, Cy5 and Cy2 recording, respectively. Four gels ran in the same conditions were analyzed (including dye swap gels) and DeCyder v6.5 software was used for the normalization and statistical analysis.

A representative gel is shown in Fig.3.9A. 17 spots were identified as differentially expressed in the stroma of wild type and *tgd4* mutant, after normalization and statistical analysis. These spots (Fig. 3.9B) were subjected to MALDI-TOF mass spectrometry, and most of the differential spots were major chloroplast proteins containing canonical transit peptides (Suppl. Table S3.1 and identification data in Dataset 1 in CD). The lack of N-glycoprotein identification in the stromal fractions encouraged us to attempt a different approach, removing major stromal contaminants and performing a N-glycoprotein enrichment step prior to a mass spectrometry study.

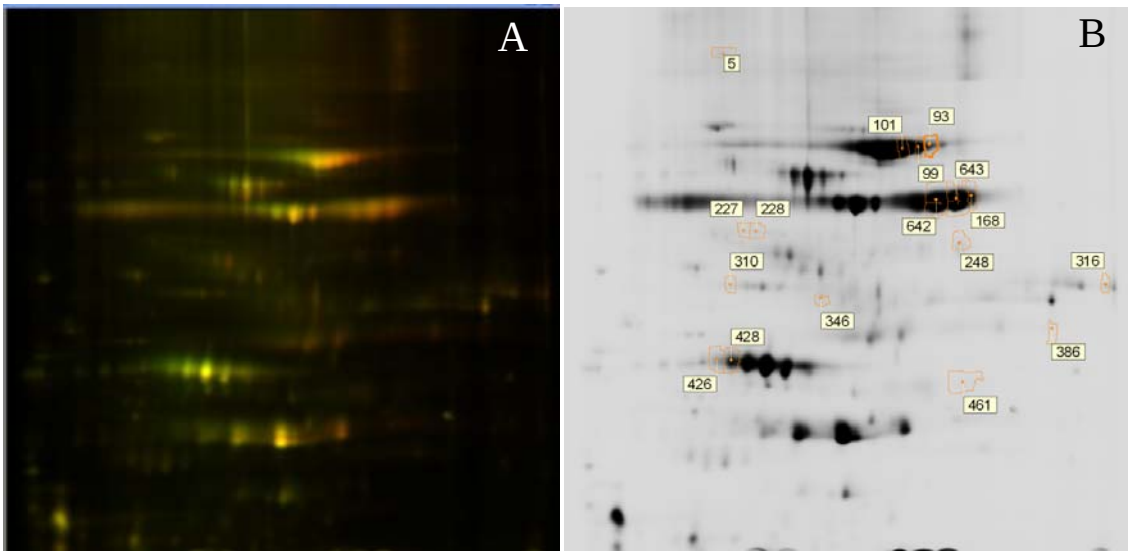

Figure 3.9: 2D-DIGE shows moderate differences between common plastid proteins, when comparing wt and *tgd4* mutant. A) Fluorescence image comparing stroma of wt and *tgd4*. cy2 (yellow), standard; cy3 (red) and cy5 (blue) 50 µg of protein per sample. B) Coomassie staining of preparatory gel. The 17 differentially expressed spots are marked

3.4.2 Depletion of major stromal proteins and high mannose N-glycoproteins enrichment

The main interest of this chapter was the identification of N-glycoproteins, preferentially decorated with high mannose N-glycans, in the chloroplast stroma, since evidence of differential protein patterns was found in this fraction. However major chloroplast soluble contaminants (i.e. Rubisco) seem to hamper the identification of minor glycoproteins. As Rubisco accounts for 40-50% of the total stroma content (Ferro *et al*., 2010), depletion of this protein in the stroma samples might improve the identification of other minor stromal proteins. Therefore, Seppro Spin Columns kit (Sigma-Aldrich) was used to obtain highly pure Rubisco free stroma fractions. Due to low protein content, 2-3 stroma samples were pooled and concentrated before Seppro purification. Subsequently, enrichment of high mannose glycoproteins based in affinity to Concanavaline A (Con A) was carried out (Suppl. Fig. S3.2). Two methods of Con A enrichment were tested, affinity columns with immobilized Con A (Pierce), and Con A-bound magnetic beads (Brucker Daltonics), being the more satisfactory results those obtained with the Con A column purification; consequently it was selected for further studies. Con A column eluted fraction exclusively contained a selected pool of high mannose type glycoproteins, as demonstrated by total digestion with Endo H (Suppl. Fig. S3.2E). A preparative acrylamide gel, stained with mass spectrometry (MS)-compatible silver staining (Blum *et al.*, 1987), was run with the rest of the purified protein sample and subjected to matrix-assisted laser desorption/ionization – time of flight (MALDI-TOF)- MS analysis for identification (section 2.6.1 of Material and Methods). Here, 6 spots were identified. Among them, 4 were typicall plastid protein, whereas the other 2 curiously presented a signal peptide: myrosinases thioglucoside glucohydrolase 1 (TGG1, AT5G26000), and thioglucoside glucohydrolase 2 (TGG2, AT5G25980) (Summary of data in Table 3.1 and details of mass spectrometry identifications in Dataset 2 in CD).

Table 3.1: Identifications in wild type stroma samples purified by Seppro and Con A columns and analyzed by MALDI TOF-MS
Summary of data of analyzed spots

	TGG1 (AT5G26000.1)	**TGG2** (AT5G25980.2)	**ATP synthase CF1 β subunit** (gi 7525040)	**Annexin arabidopsis 1** (AT1G35720)	**Fructose biphospahte aldolase** (gi 8399660)	**Ferredoxin NADP (+) oxidorreductase 2** (AT1G20020)
Length(aa)	541	536	248	317	399	255
Nominal mass (M_r)	61132	61885	53900	36296	43075	28903
pI	6,04	7.11	5,38	5.21	6,18	5.74
Score	71	702.5	36	72	68	73
Coverage (%)	14	17	3	7	21	37
Peptides matching	6	9	1	1	7	1
Predicted Localization (Target P)	Secreted (SP: 0,963)	Secreted (SP: 0,895)	Chloroplast (N.d (Other: 0,234))	Chloroplast (N.d (Other: 0,905))	Chloroplast (cTP: 0,758)	Chloroplast (N.d (Other: 0,903))
No. Predicted N-Glyco sequon (netNGlyco)	5	2	1	1	0	1

These proteins are described to harbor high mannose N-glycans (Zhou *et al.*, 2012; Kim *et al.*, 2013; Liebminger *et al.*, 2012) and are predicted to have SP. Confirmation of the presence of high mannose type N-glycans was visualized after digestion with Endo H as a molecular mass shift, with TGG1 antibody, kindly provided by J. Meijer from Uppsala University (Fig 3.10). This experiment confirmed the high mannose type nature, in accordance with previously mentioned publications. Additionally, distribution of TGG1 and different markers in the microsome fraction was analysed by sucrose gradients (Suppl. Fig S3.3A). TGG1 signal mainly co-localized with that of the ER marker BiP, as expected for a signal peptide carrying protein. Co-localization of TGG1 with Rubisco (chloroplast stroma marker) was also found. Unfortunately, although experiments were repeated several times, no conclusive results can be raised about this specific fact, since Rubisco was present in almost all fractions (Suppl. Fig. S3.3B). Although some groups have reported TGGs localization in plasma membrane by fluorescent fusions of the proteis (Agee *et al.*, 2010), it would be necessary to carry out these experiments with our specifically purified samples in order to check this fact. Therefore, an additional approach was attempted, in this case trying to highlight the presence of minor-abundant proteins within a complex mixture. For that purpose, chloroplast samples were subjected to purification with a specific kit (ProteoMiner, BioRad) that normalizes the presence of low-abundant proteins in a complex mixture (section 2.5.3 of Material and Methods). It was tested on wild type total chloroplasts fractions, and the purified protein extract was run in a SDS-PAGE gel, stained with Coomassie colloidal and subsequently analyzed with a high accurate technique, LC-MS (LC-LTQ Orbitrap). However, only major highly abundant chloroplast proteins were identified, and consequently this approach was discarded for further experiments (Suppl. Table S3.2 in Dataset 3 in CD).

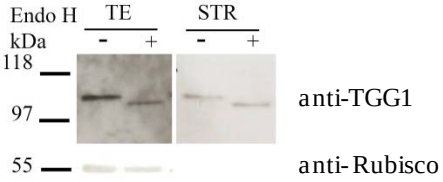

Figure 3.10: Myrosinase TGG1 harbours high mannose N-glycans. Endo H digestion of TE and purified STR samples by Seppro and Con A columns, bloted against TGG1 and Rubisco antibodies. 5 µg of protein per line. TE, total leaf extract, STR, stroma.

Given that none of the previous techniques seemed to be sensitive enough for detection of low abundant proteins, a different method should be attempted, involving a more sensitive measurement, capable of accurately detect minor or low-abundant high mannose type glycoproteins. Additionally, sample obtention and preparation is challenging. Thus, although identification of myrosinases seemed promising, it would be useful to find another technique which allow less handling of the samples, the use of less amount of material, and being more sensitive in order to identify proteins of interest in a more direct and quick way. Additionally it would be interesting to identify TGGs from another approach, and to check whether there can be found more proteins with similar characteristics. Then, it was decided to analyze samples with the

nowadays more accurate technique, tandem LC-MS (LC-MS/MS), but following a procedure markedly more focused in identification of N-glycoproteins

3.4.3 Liquid chromatography coupled to tandem mass spectrometry (LC-MS/MS) focused in detection of N-glycoproteins

As briefly mentioned before, LC-MS is a highly sensitive technique used for accurate identification of proteins in complex mixtures (Mallick and Kuster, 2010). In this case, samples were analyzed by MS without any previous separation in gel, in order to minimize the loss of material.

Analysis of purity and quality of total chloroplast and stroma preparations for MS experiments:

Wild type and *tgd4* total chloroplast preparations were previously subjected to immunoblot analysis with several markers. Anti-β(1,2) xylose antibody was used for the detection of proteins harboring complex N-glycans, ConA peroxidase lectin affinity was used for detection of proteins harboring high mannose N- glycans. (Fig. 3.11). According to our empiric experience, samples were considered as suitable for MS analyses.

Nevertheless, a previous test in LC-MS/MS was performed before specific spectrometric assays, in order to have a previous overview of their purity and complexity. For that, total chloroplast samples were subjected to Filter Aided Sample Preparation (FASP) and analyzed by LC-MS/MS (LC-LTQ-Orbitrap). FASP is a technique that allows sample preparation with little manipulation, as all steps are carried out in a filter device. See section 2.6.2 of Material and Methods for detailed description. Samples are treated with a series of buffers and digested with trypsin, obtaining a mixture of peptides which are further analyzed by LC-MS/MS. The obtained dataset of identified peptides was analyzed using OMSSA algorithm (Geer *et al.,* 2004), and proteome composition and predicted subcellular localization were manually determined based in criteria suggested by van Wijk and Baginsky, 2011 (Fig. 3.12A). It was confirmed that the samples were highly pure (less than 5% contaminant peptides) in both cases, which

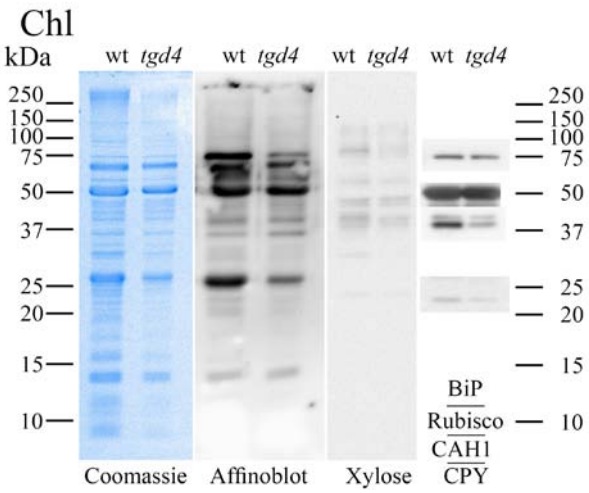

Figure 3.11: Control of wild type and *tgd4* chloroplast preparations. Total chloroplast preparations subjected to Coomassie blue staining ConA lectin affinoblot and immunoblot assays using anti-xylose, antiBiP, anti-CPY, anti-Rubisco and anti-CAH1 antibodies.Chl, chloroplasts; wt, wild type.

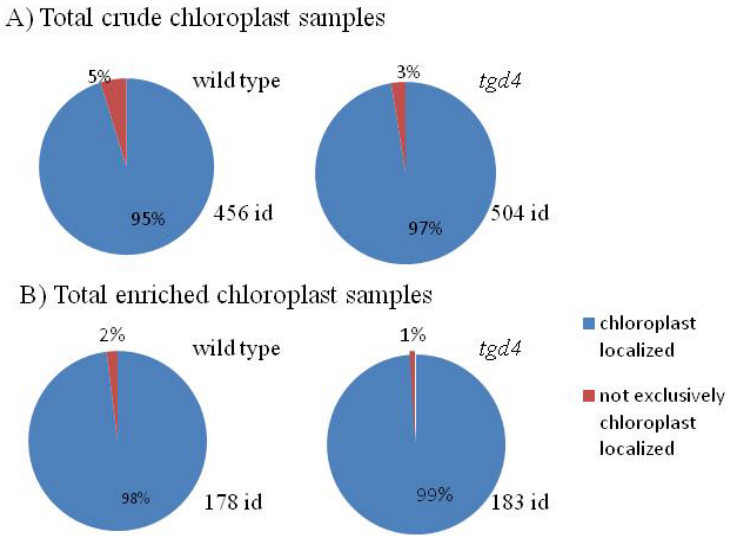

Figure 3.12: Analysis of chloroplast proteome reveals samples are highly pure. A) Total crude chloroplast samples. **Left panel:** wild type sample, 456 proteins were identified, 95% of them were predicted to locate mainly in chloroplast, 4,82% were predicted to be located totally or mainly in other parts of the cell. Considering spectral counts (SCs), an indicative of protein abundance, the abundance of these minority proteins is extremely low, between $4{,}95.10^{-4}$% (1 SC) - $9{,}15.10^{-4}$% (2 SCs) of relative frequence**. Right panel:** *tgd4* sample*,* 504 proteins were identified. 2,57% were predicted to be located totally or mainly in other parts of the cell. Considering SCs, the abundance of these non-exclusively-plastid proteins is around $4{,}37.10^{-4}$% (1 SC) and $8{,}74.10^{-4}$% (2SCs). **B)** Total N-glycoprotein-enriched chloroplast samples. **Left panel:** wild type, 178 proteins were identified. 97,8% of them were predicted to locate mainly in chloroplast. Proteins that were predicted to be located totally or mainly in other parts of the cell: 2,2%. Considering SCs, the abundance of these non-exclusively-plastid proteins is around $4{,}25.10^{-4}$% (1 SC) and $9{,}07.10^{-4}$% (2SCs). **Right panel:** *tgd4,* 183 proteins were identified. 98,9% of identified proteins were predicted to locate mainly in chloroplast. Proteins predicted to be located totally or mainly in other parts of the cell: 1,1%. Considering SCs, the abundance of these non-exclusively-plastid proteins is around 5.10^{-4}% (1 SC) and $8{,}93.10^{-4}$% (2SCs) **CRITERIA**: Localization of each protein identified by LC-MS/MS after filtration of data with OMSSA algorithm, was checked manually under different online databases: SUBA3, PPDB and TAIR, as well as the other annotations in predictors like TargetP, amiGO, etc, indexed in these sites. Proteins in which most of the consulted database agreed in chloroplast localization, were considered as "mainly located in chloroplasts", while proteins in which were not a strong agreement or even any reference to plastid localization, were considered "non located/mainly not located in chloroplast". Id, idenfitication.

is consistent with no significant differences in proteome composition previously observed in immunoblots (see Suppl. Tables S3.3 andS3.4 in Dataset 4, in CD).

In addition, general proteome of wild type and *tgd4 stroma* samples was also analyzed. The proteome of this fraction was expected to contain fewer proteins, being much simpler than the total chloroplasts including membrane proteins. Indeed, identifications performed by LC-MS/MS were approximately half of those in total chloroplast samples, and proteins were identified with lower spectral counts (see Suppl. Tables S3.5 and S3.6). In this case, also nearly all identified proteins were chloroplast localized, and, like total chloroplast, the number of identifications was similar for wild type and *tgd4* stroma fractions. This fact, together with the decrease of protein amounts that would take place after enrichment with Con A, would lead to extremely low protein amount that would impede significant detection by MS. These limitation lead to the decision of analyze only total chloroplast fractions.

Analysis of N-glyco-enriched chloroplast samples

In order to increase the relative amount of glycoproteins, specific FASP method was carried out, in which subsequent steps of binding of the digested peptides to Con A, and digestion with PNGase F, a glycosidase that releases the link of the asparagine to the N-glycan- were performed, prior to LC-MS/MS analysis (section 2.5.5 of Material and Methods). The action of this enzyme produces a change in the asparagine structure, which becomes an aspartic acid, and consequently peptide mass is incremented in 0,984 Da. This shift on mass can be detected by the mass spectrometer. In some biological replicates, N-glyco FASP was performed in the presence of labeled water ($H_2{}^{18}O$) so glycosylated asparagine residues are converted into ^{18}O-labeled aspartic acids, provoking an increase of mass of 2,989 Da, and then increasing the confidence of identifications (Kunster and Mann, 1999).

After Con A enrichment of total chloroplasts, the number of protein identifications was about 2,5 fold lower than the non-enriched samples (Fig. 3.12B). A major amount of Rubisco is still present in wild type and *tgd4* samples, although it is 4 fold lower than crude samples. The specificity of the purification method has been previously reported (Zielinska *et al.,* 2010), therefore, this fact would point to co-purification of this protein with N-glycoproteins due to protein-protein interaction. Nevertheless, as it can be observed in Fig. 3.12, after purification samples maintains their high purity, as crude samples, in accordance with the specific purification performed, and demonstrating that an important number of proteins have been eluted from the sample. In general terms, not striking differences between wt and *tgd4*, neither according to presence/absence of proteins nor related to number of spectral counts, were observed. Lack of significant differences indicates that proteins of interest are present in the chloroplast in a very low amount.

Identification of N-glycoproteins:

Once the enriched samples were obtained (section 2.6.3 of Material and Methods) and their quality checked, resulting data was analyzed with three different algorithms; OMSSA (Geer *et al.,* 2004), X!Tandem (Craig and Beavis, 2004), and SEQUEST (Eng *et al.,* 1994) (section 2.6.3 of Material and Methods). The more interesting identified proteins are listed in Table 3.2. Most of them have been previously described as plastid proteins containing a transit peptide, and some of them, i.e myrosynases, were also found in our previous proteomic assays (Table 3.1). After Con A purification, proteins not found in previous LC-MS/MS of not enriched (i.e, crude) total chloroplast samples were identified, for example Rubisco Methlytransferase (AT1G14030.1), a protein member of RING U-box superfamily (AT4G05350.1), cytochrome P450, CY708A3 (AT1G78490.1), sucrose synthase 3, SUS3 (AT4G02280.1) and a protein from the family Zinc finger CCCH-type (AT1G32360.1). However, for the purposes of this study, the more interesting identifications were myrosinases TGG1 (AT5G26000.1) and TGG2 (AT5G25980.2), which appeared again in the samples (Table 3.2 and 3.3 and Suppl. Tables S3.7 to S3.29 in Dataset 4).

Table 3.2 Summary of identified N-glycoproteins						
Prot identified	Chlwt 030412	Chltwt 290312	Chltwt 040412	Chl *tgd4* 110412	Chl *tgd4* 041111	Chl *tgd4* 120412
TGG1 (AT5G26000.1)	(S/O/X!)	(S/O/X!)	(S/O/X!)	(S/O/X!)	(S/O/X!)	(O/X!)(O*X!*)
ROC 4 (AT3G62030.1)	(O/X!)	(S/O/X!)(X!*)	(S/O/X!)	(S/O)	(S/O)(X!*)	(O/X!)
TGG2 (AT5G25980.2)	(S/O/X!)	(S/O/X!)	(S/O/X!)	-	(S/O/X!)(X!*)	(O(X!)
PRK (AT1G32060.1)	(S/O/X!)	(S/O/X!)(X!*)	(S/O/X!)	-	(O/X!)(X!*)	(O)
PSBR (AT1G79040.1)	-	(S/O/X!)	(S/O/X!)	(S/O/X!)	(S/O/X!)(X!*)	(X!)
CaS (AT5G23060.1)	(S/O/X!)	(S/O/X!)	-	-	-	-
RNAbinpr (AT4G24770.1)	-	(S)	(S)	-	-	-
FTSH2 (AT2G30950.1)/8 (AT1G06430.1)	(S/O/X!)	-	-	-	-	-
RubMettr (AT1G14030.1)	.	-	(S)	-	-	-
RING/Ubox (AT4G05350.1)	-	(S)	-	-	-	-
CY708A3 (AT1G78490.1)	-	-	(S)	-	-	-
SUS3 (AT4G02280.1)	-	-	(S)	-	-	-
Zinc finger (CCCH-type) family protein (AT1G32360.1)	-	-	-	-	-	(O)

Table 3.2: Identified N-glycoproteins. Summary of the most reliable proteins identified in chloroplast samples of wild type and *tgd4* mutant. Complete lists of identifications were analyzed with three different algorithms: S (SEQUEST); O (OMSSA); X! (X!Tandem). Asterisk (*) indicates N-glycosylation site identified in presence of $H_2{}^{18}O$ water (only performed in Chl wt 030412, Chl wt 290312, Chl tgd4 041111, Chl tgd4 1204012 samples). Chl, chloroplasts; wt, wild type.

Additionally to these assays performed, alternative approaches also based in LC-MS/MS were carried out, in order to check whether another interesting identification was achieved. On one hand, one wild type and one *tgd4* chloroplasts preparations were subjected to *in source fragmentation* (ISF), a technique which provides more accurate information focusing on glycan structure. As a result, a disease resistance protein RPP1-like R1(RPP1-like) (gi 215261582) was identified only in the wild type sample. On the other hand, knowing that PNGase F is unable to cleave N-glycans harboring fucose residues in α(1,3) position, N-glyco FASP using PNGase A in the digestion step was performed. Experiment was performed in two wt biological replicates, and the obtained data were analyzed under X!Tandem algorithm. As a result, it was identified Photosystem II subunit R, PSBR, (AT1G79040.1), that also appeared in the previous analysis (Suppl Table S3.30). However, as these results were not as consistent as those previously performed (Table 3.2), thus finally they were not considered.

Table 3.3 Representative data from identifications performed by LC-MS/MS and processed with X!Tandem from samples Chl wt 290312 and Chl *tgd4* 041111

Prot identified / Peptide	Theoric mass	Charge		E-value		Prediction of protein location (Target P)	Number of predicted N-glycosylation sites(NetNglyc)
		wt	*tgd4*	wt	*tgd4*		
TGG1 (AT5G26000.1)							
DIDVmDELnSTGYR	16447062	2	2	6.7e-012	2.1e-011	Secreted (SP: 0,963)	5
FGLSYVDFAnITGDR	16757967	2		7.5e-012			
nATGHAPGPPFNAASYYYPK	21239826	3	3	1.1e-005	3.7e-007		
ROC 4 (AT3G62030.1)							
DFmIQGGDFTEGnGTGGISIYGAK	24520977	2	2	5.6e-006	3.6e-011	Chloroplast (cTP:0.927)	4
TGG2 (AT5G25980.2)							
FGLSYVDFNnVTADR	17188025	2	3	3.1e-008	5.8e-010	Secreted (SP: 0,895	2
LTAmTDSLAnLTSLDANGQPPGPPFSK	27603401	3	3	2.1e-002	3.8e-004		
PRK (AT1G32060.1)							
HADFPGSNnGTGLFQTIVGLK	21740881	3	3	5.0e-009	5.9e-009	Chloroplast (cTP:0.739)	2
PSBR (AT1G79040.1)							
TDKPFGInGSmDLR	15677425	2	3	3.4e-003	5.0e-004	Chloroplast (cTP:0.634)	1
CaS (AT5G23060.1)							
LGTDSYnFSFAQVLSPSR	19899557	2		1.9e-005		Chloroplast (cTP:0.934)	1

Thus, summarizing, the identification of myrosinases TGG1 and TGG2 is noticeable. Interestingly, they were also previously identified in stroma purified samples (Table 3.1). Suppl Fig.S3.4 shows a representative diagram of mass spectrometry obtained during analyses. Most of the identified peptides were found to belong to transit peptide containing proteins, similar to those found in preliminary analysis (Table 3.3). Interestingly, myrosinases TGG1 and TGG2 appeared in all replicates after data analysis with the three named algorithms, and after performing the experiment in presence of labeled water, giving more confidence to the identifications. At the same time, some details about their sequence are really interesting: They contain a longer C terminal than other TGGs and, as CAH1, they do not have TP but SP as previously shown. Although myrosinases are described as vacuole, apoplast resident proteins (according to TAIR) and peroxisomes (Reumann *et al.*, 2007), involved in plant defense (Barth and Jander, 2006), it is also true that they have been located in chloroplast in different proteomic studies (Zybalov *et al*., 2008; Kleffmann *et al*., 2004; Simm *et al*., 2013).

Then, it would be really interesting to deeply study these myrosinases, and the underlying mechanisms they use for their transport, taking into account the alternative possibilities they could take, suggested by this study or literature (see Discussion).

IX. DISCUSSION

Arabidopsis thaliana α carbonic anhydrase 1(CAH1), has been described as the first example of an N-glycoprotein targeted to the chloroplast through the endomembrane system in dicot plants (Villarejo *et al.*, 2005). Little was known about the specific influence of potential N-glycosylation sites and disulphide bonds over the protein folding and function. In addition, its carboxyl terminal sequence (C terminus) seems to present special features (Villarejo *et al.*, 2005) that might affect its transport, although it remained uncharacterized. In order to unravel the significance of post translational modifications in *Arabidopsis* CAH1 stability and function, an hemmaglutinine (HA) epitope tagged version was used (Fig. 1.1). HA epitope tag fusion hardly alters the structure of the proteins (Zhao *et al.*, 2013) decreasing the risk of misfolding and/or mistargeting when compared to fluorescent proteins fusions. In the present case, the protein pattern of HA tagged CAH1 (HC) was similar to the native protein, and therefore was considered suitable for the analyses (see Chapter 1).

N-glycosylation sites were successfully mutated, and partially or non-glycosylated (NG) versions of HC were detected (Fig. 1.3). The amount of protein harboring high mannose N-glycans increment in *Arabidopsis* protoplasts and *Nicotiana* plants transiently expressing NG (Fig. 1.5), suggesting an increased retention in ER. This would strengthen the already reported influence of N-glycosylation presence/absence in proper protein folding. However, there have been reported particular cases in which elimination of all or some of the glycans has moderate or no effect on stability or folding (Parodi, 2000; Mitra *et al.*, 2006). In this line, other studies concluded that the effect of N-glycan removal on protein stability depends on the site mutated and varies widely among different proteins (Pagny *et al.*, 2003; Shental-Bechor and Levy, 2008). Taking this into account, it could be the case of the CAH1's N-glycosylation site N4, which seems to be partially glycosylated (Fig. 1.5). N4 is spatially close to the cysteine residues C1 and C3 (Fig. 1.2). These cysteines seem to form a disulfide bond, which is required for normal protein folding and function (Burén *et al.*, 2011). Therefore, the formation of such bond seems to hamper the occupation of N4 glycosylation site, suggesting that N4 glycan acquisition is probably not essential for CAH1 stability and/or folding. Nevertheless, the lack of all N-glycans in the non-glycosylated CAH1 (NG) might cause a general alteration (destabilization) of its three dimensional structure, likely producing ER retention of the altered protein. This hypothesis is strongly supported by the noticeably increase in chaperone binding to the NG form (Fig. 1.6).

In a similar way, disulphide bond disruption, which is known to affect the protein general structure (Betz, 1993), also produced CAH1 harboring nearly only high mannose type N-glycans (although lower signal intensity was detected in western blots analysis, C1C3 was highly sensitive to Endo H treatments) (Fig. 1.9). Unfortunately, it was not possible to clone the C1C3

construction into binary vector to express it in *Nicotiana* leaves, so it could only be expressed in *Arabidopsis* protoplasts, which yielded no appreciable signal in BiP co-precipitation assay due to limitation of plant material. Therefore, it is not possible to indubitably conclude the reason why C1C3 mutant increases the level of high mannose type N-glycans when compared to HC, whether it is due to a general protein misfolding and degradation, or another reason. However, it is noticeable that C1C3 mutation seems to decrease protein expression levels (Fig. 1.9A), a fact previously seen in cell cultured protoplasts (Burén *et al.*, 2011). Additionally, this work reported an increase in chaperone binding in *Arabidopsis* cell culture protoplasts when either C1 or C3 were mutated, hampering the formation of the mentioned disulfide bond (Fig. 1.8). This fact supports the hypothesis that a general protein structure alteration or misfolding might be also the case in transiently transformed *Nicotiana* leaves. Summarizing, post-translational modifications clearly influence CAH1 folding, trafficking and function, having particularly N-glycosylations a pronounced effect in protein folding.

Protein C terminus has been described as relevant in different situations and proteins, like its requirement for protein activity in the xylosyltransferase of *Arabidopsis thaliana* (Pagny *et al.* 2003). The C terminus of CAH1differs from those of other plant αCAs (Villarejo *et al.,* 2005). It is highly hydrophilic and enriched in lysine residues, similar to the transit peptide of apicomplexan proteins, which are transported through several membranes of the complex plastid before reaching the stromal compartment (see section 1 of Introduction). Indeed, highly hydrophilic C terminal domains containing several lysine residues (as CAH1) are known to be involved in protein stability (Wong and Ho, 2013). Since the function of CAH1 C terminal was still unknown, its involvement in CAH1 stability and trafficking was studied. Mutation of two of the last four lysines in the CAH1 C terminus (Cmut) produced a heterogeneous glycosylation pattern, with isoforms presenting both, high mannose and complex N-glycans, especially noticeably in the *Nicotiana* samples and comparable to HC (Fig. 1.10). Indeed, the low BiP binding to Cmut corroborates that this point mutation in the C terminal does not induce the unfolded protein response in the ER, excluding the involvement of the C terminal polylysine sequence in protein stability/folding.

On the other hand, removal of part of the C terminus including the polylysine sequences (Cdel), increased the sensitivity to Endo H digestion, due to increased high mannose type N-glycans when compared to HC (Fig. 1.10). This fact suggests that the protein is partially retained in the ER. Intriguingly, moderate levels of BiP binding (Fig. 1.11) could indicate that despite the lack of native N-glycosylation, the protein is properly folded. In this line, there are evidences demonstrating that the immediate three sugar residues closest to the protein are sufficient for stabilization (Price *et al.,* 2010). In addition, different studies in plants have reported that under normal growth conditions, the type of N-glycosylation, regardless it is complex or high manosse, is not decisive for the function of the protein (von Schaewen *et al.*, 1993; Liebminger *et al.*,

2013). This could mean that although not harboring native N-glycans, the protein is folded and consequently, probably functional. In this line, this can be reason of absence of indicative changes in the expression levels of KOR1 (Fig. 2.4), regardless it was properly or defective glycosylated. Although the underlying mechanisms that determine whether a site is N-glycosylated or not or the type of N-glycan attached are not completely clear yet (Ruiz-May, 2012), several possibilities can be depicted in the HC mutants. Particularly in the case of deletion of part of C terminus, the more surprising result, the sort of N-glycans acquired by Cdel construction could be determined by a combination of factors. Maybe is an effect caused by its overexpression Another possibility is that truncated proteins might not reach the final destination due to general misfolding. Certainly, there are evidences reporting that sugar residues can influence on the folding kinetics, causing ER retention by the quality control system (Mitra *et al.*, 2003; Herbert and Molinari, 2007). Lastly, the three dimensional structure of Cdel construction might be altered, somehow hampering the access of Golgi glycosyltransferases to the protein. This would lead to correctly folded proteins with increased amount of high mannose N-glycans compared to the wild type. Indeed, previous reports on mammal models have shown proteins harboring high mannose N-glycans, typical from the ER, being transported through the Golgi apparatus without further modifications (Stanley, 2011). The strong requirement of posttranslational modifications for CAH1 proper folding and function found in our studies, together with suggested similarities of its C terminus to N-glycoproteins from apicomplexa, support the possible maintenance of endomembrane route to chloroplast over evolution (Villarejo *et al.*, 2005; Bhattacharya *et al.*, 2007; Cavalier-Smith, 2008). However, a recent phylogenetic study has concluded that this pathway is not a relic or ancient one, since the N-glycoproteins known to date do not have prokaryotic but eukaryotic origin (Gagat *et al*, 2013). Nevertheless, more studies about this item, from different approaches, should be carried out in order to purpose a clear explanation. Hopefully, development of *Arabidopsis thaliana* CAH1 knock-out lines stably transformed with Cmut, Cdel and NG mutants would shed light on the requirement of post-translational modifications for protein function.

Function of complex N-glycans, at a more general level, is not deeply characterized in plants, and, until now, scarce information about their biological significance is available (see section 3.1.1 of Introduction). With the aim of studying the role of different sugar residues in tolerance to abiotic stress of plants, an analysis from a physiological approach using *Arabidopsis* mutants impaired in proper Golgi N-glycosylation activity was performed. The aim of the present work was to check whether our results were analogous to those described in Kang *et al.*(2008), and additionally, if there were an specific role of fucoses added by *Arabidopsis* core α(1,3)fucosyltransferase (FT). For that, different analysis under salt/osmotic stress were performed in two mutants whose N-glycosylation pathway is disrupted at some step in Golgi: double knockout of α(1,3)FTs, ΔFT, produced by Prof. Samuelssons group, and GnTI knockout, *cgl1*. As previously mentioned, the

function of complex N-glycans in plants is less characterized than in other organisms, however, in salt stress conditions, a general trend pointing to higher arrest of root growth, proportional to more extended N-glycosylation alteration, was observed (Fig. 2.1). Results are consistent with those obtained in Kang *et al.*, 2008, where root growth in single mutant XylT and other double mutant FT were less affected than *cgl1*, whose proteins lacks α(1,3)fucose and β(1,2)xylose residues, and triple mutant *xylT-fucT a-fucT b*, which lacks both xylosyl- and fucosyl transferases. This fact suggested that fucose and xylose residues apparently display overlapping functions, as mutants harbouring only α(1,3)fucose or only β(1,2)xylose were not as affected as mutants with glycoproteins lacking both residues (Kang *et al.*, 2008). This is in accordance with results obtained in this study (Fig. 2.1), and those reported in Strasser *et al.*, 2006. Additionally, Kang *et al.*, 2008 and Strasser *et al.*, 2006 point to a potential independence of the xylosylation and fucosylation pathways. Our results, although only in an intuitive way, show a sort of intermediate phenotype in the mutant that lack fucose (ΔFT), when compared with wild type and *cgl1*. This could be considered as an indirect evidence supporting this idea of the independence of the pathways, meaning that plants are able to overcome, somehow, the absence of β(1,2)xylose or α(1,3)fucose in its glycoproteins, although in a lesser extent than in wild type conditions.

Once checked the affection in root growth, it was decided to study another physiological aspect that needs a proper N-glycosylation for being carried out, that is, generation of a consistent cell wall (see section 2.2 in Chapter 2). In line with the results analyzing root growth, a similar trend was observed when treating leaf tissue with cell wall degrading enzymes (Fig. 2.2), as the mutant more affected in N-glycosylation, *cgl1*, seems to be the more sensitive to degradation, followed by ΔFT. However, high variability of the data does not allow formulate any definite conclusion. Compiling all results, it seems that lack of proper complex N-glycosylation under salt or osmotic stress probably leads to proteins that are not fully functional at some extent. It seems that proteins presenting high mannose or non-α(1,3)fucosylated glycans may be not completely functional, at least in conditions of abiotic stress. Taking the results together, it could be suggested that high mannose glycans structures are *per se* especially sensitive to abiotic stress. In this line, Kang *et al.*, 2008 suggests that xylose and fucose moieties could play a role in some kind of protection of proteins against hostile agents in the media. Maybe, complex N-glycans protect proteins by provoking steric impediments that hampers structural damage. This would allow the protein to be as highly functional as possible under stress conditions as may happen in wild type plants. It would be interesting to check whether high mannose N-glycans structures confer similar sensitiveness to proteins under *biotic* stress situations, as it would reinforce the hypothesis of the role of complex glycans in protein folding assistance.

In relation to cell wall mechanical strength, it has been reported that proper N-glycosylation is required for proper cellulose biosynthesis (see Chapter 2). This could be related with the special

susceptibility of the cell wall to salt stress conditions observed in defective N-glycosylation mutants. Additionally, and being consistent with the observed results, a reduced, affected or impaired cellulose biosynthesis –due to a defective glycosylation- can lead to a weaker cell wall, but clear effects would be only observed under salt stress conditions.

Regarding high variability of data obtained, and in order to obtain more significant evidences about cell wall alteration under abiotic/osmotic stress conditions, repetition of some experiments would be needed, or even the modification of some experimental design aspects. In this line, variability observed in experiments involving protoplasts is in accordance to that observed in transient expression experiments, confirming once more the delicacy of this biological system. Indeed, experiments evaluating cell wall growth were not conclusive, since its formation could not be measured despite the different approaches and culture media used. Actually, estimation of cell wall growth in similar conditions has been performed by Kwon *et al.*, (2005), being the most reliable indication of such formation electronic microscopy assays. Although Kwon's study observed calcofluor fluorescence only after 1h in *Arabidopsis* protoplasts, other groups could not detect cell wall formation within the same time period (Kohorn *et al.,* 2006). Kohorn and co-workers reported calcofluor staining only after 72 h of *Arabidopsis* protoplast culture, a similar time lapse described for cell wall formation in cotton cells protoplast (Yang *et al.* 2008). In our study the absence of calcofluor fluorescence was consistent in both, plate reader and microscopy experimental designs in which fluorescence was measured in even shorter periods. Considering the publications previously mentioned, the experimental design might be improved just by extending the experimental time period. However, a similar approach by Abreu and coworkers could not detect calcofluor fluorescence even after seven days of healthy protoplast culture (personal communication). Therefore, a different strategy maybe focused on scanning electron microscopy should be attempted, in order to achieve cell wall growth measurements.

As previously mentioned, the involvement of the N-glycoprotein endo β(1,4)glucanase (KOR1) in cellulose biosynthesis, has been reported, and alterations in its N-glycosylation state, together with osmotic stress decreased plant performance producing a characteristic phenotype (Kang *et al.*, 2008). As previously mentioned, the importance of N-glycosylation in enzymes involved in cellulose biosynthesis, like KOR1, has been reported (Lukowitz *et al.,* 2001). Mutations in KOR1 resulted in a series of strong and varied affections, leading to the conclusion that this protein is involved in the mentioned cellulose synthesis of primary and secondary cell walls, xylem vessel development, pectin composition, and cytoskeletal organization (His *et al.,* 2001; Lane *et al.,* 2001; Sato *et al.*, 2001; Szyjanowicz *et al.,* 2004; Paredez *et al.,* 2008). Also is involved in release of newly synthesized cellulose microfibrils and alteration of the crystalline properties of microfibrils (Szyjanowicz *et al.,* 2004, Takahashi *et al.,* 2009). KOR1 N-glycosylation state has

been shown to influence cell wall functionality under salt stress (Kang *et al.*, 2008). For that, expression levels of the N-glycoprotein KOR1 were analyzed. The observed high variability of the data impeded a definite conclusion that might indicate differences between wild type and N-glycosylation mutants (Fig 2.4). Not statistical significance of KOR1 expression levels could be due to intrinsic insensitivity of the protein to the type of N-glycosylation its harboring, neither the stress conditions, or maybe also influences the high sensitivity of the RT-PCR technique, in which obtained variations lead to not statistical significant data. These results outline the difficulty of the experiments and the necessity of provide a new focus for carry on the study. In order to overcome the problem, some options would be to repeat the experiment more times in more biological replicates, till some statistically significant tendency would be observed, or perform additional experimental assays, like Northern-blot or quantitative PCR (qPCR). It also could be tried a proteomic approach, like extraction and measurement of the protein by different methods. Furthermore, it would be interesting to test additional stresses in order to find more relations with N-glycosylation and identify the residues exactly involved in protection of proteins, and rule out how this mechanism operates.

The exact localization of the enzymes involved in the process of late N-glycosylation is still controversial, as previously mentioned (see also section 4.4 of Introduction). In spite of the moderate, although no statistical significant, tendency observed pointing to a potential partial role of fucose residues in stability of N-glycoproteins under salt stress conditions, studying α(1,3) fucosylation is still noteworthy. Its immunogenic nature together with the fact that, up to date, α(1,3) fucose is the only residue in the N-glycosylation pathway known to be added by two different enzymes, converts FTs in an interesting issue for analysis. A potential redundant function of FT11 and FT12 was suggested by Strasser *et al.*, 2006, that would guarantee proper fucosylation. However, Forth *et al.*, (unpublished), suggested a predominant activity of FT12, specifically in a pool of proteins targeted, particularly, to the plastid. In that case, FTs would fucosylate different sets of proteins, could be possible that they would localize in different Golgi stacks. Therefore, in the present work, with the aim to shed light on this issue, an attempt to characterize their localization was performed. Ruling out whether FTs are located in the same Golgi stacks or they are in different (and specialized) ones, would imply the production of proteins targeted to a specific subcellular destination. Up to date, none of the proposed models in different studies can explain, in general, the sub Golgi distribution of a given N-glycan processing enzyme in plant cells, leading to the conclusion that signal and mechanisms for accurate localization in Golgi have to be determined for each enzyme individually. As explained in sections 4.4 of Introduction and Chapter 2, the way Golgi glycosyltransferases arrange into stacks to carry out N-glycosylation still is not fully understood, as is no clear whether enzymes are fixed or whether they can act at early stages (Bencúr *et al.*, 2005; Strasser *et al.*, 2006; Kajiura *et al.*, 2012), or there are fixed but maintained by a system of vesicles (Tu and Banfield, 2010). In

animal models, this issue has been extensively studied, and different studies support the steady-state distribution (Rhee *et al.*, 2005).

Evidence of Golgi localization of α(1,3)FTs had been described by confocal microscopy (Forth *et al.*, unpublished). However, the aim of this work was to discern whether both FT11 and FT12 were in the same or different Golgi stacks. Therefore, accurate localization was attempted by electron microscopy (TEM). For that, cryo sections of two and four weeks old *Arabidopsis* leaves of plants expressing both tagged FTs, were subjected to TEM analysis (Fig. 2.8).

Concurrently, a complementary analysis performed in isolated microsome showed enrichment of tagged FTs in this fraction, which includes Golgi cisternae. However, hardly detection could be achieved by specific antibodies, suggesting that must be very few FT11-myc and FT12-HA tagged fucosyltransferases within the Golgi apparatus (Fig. 2.7). To our knowledge, the present work is the first attempt of subcellular localization of both *Arabidopsis* FT isoforms. However, TEM immunogold analysis of FT11 and FT12 could not provide definite information of the specific subcellular location, due to low labeling and scarce gold particle number (Fig. 2.8). Increase of primary and/or secondary antibody concentration lead to high background signal, hampering specific detection of the FTs (not shown). The punctuated distribution of Golgi cisternae together with its small size compared to the total cell volume hampers its detection in ultrathin sections, as stated by Han *et al.*, (2013). The low labeling in the present experiments also reflect the difficulty of immunogold localization of minor Golgi resident proteins. What is more, TEM analysis of any aspect of Golgi apparatus is still a challenge, as current techniques present several limitations for dealing with a highly dynamic organelle, and not very abundant structure within the cell. Therefore, development of different subcellular visualization methods or techniques would be beneficial for accurate detection and localization of FTs.

Exploration of plastid N-glycoproteome is currently developing, and advances remains to be achieved, being technical limitations one of the barriers still to be overcome (van Wijk and Baginsky, 2011). Several proteomic analyses have reported signs of occurrence of similar proteins to *Arabidopsis* CAH1 in chloroplast, opening a door to new protein identifications. Consequently, new insights in their biological function and possible use in biotechnological applications should be attempted. Therefore, a study of plant N-glycosylation focused in the identification of chloroplast glycoproteins, was performed.

The existence of more *Arabidopsis* plastid N-glycoproteins, was proposed by Villarejo *et al.*, 2005. Additionally, clear evidence of proteins harboring high mannose N-glycans (HMGP), typical from ER in *Arabidopsis* chloroplast were obtained in our laboratory (Fig. 3.6, 3.7 and 3.8) and by other research groups (Nanjo *et al.*, 2006). The presence of HMGP lacking further modification suggest the possibility of a potential non-described transport mechanism,

independent of other chloroplasts protein pathways, in which high mannose N-glycoproteins would be imported directly from the ER through the chloroplast envelope.

Regarding this fact, there is controversy in scientific community when defining how the traffic of proteins harboring different types of N-glycans takes place. The presence, in different parts of the cell, of proteins harboring N-glycans typically acquired in ER (that is, high mannose), has been extensively discussed. Some authors suggest that some proteins can pass through Golgi apparatus without being modified by Golgi resident glycosyltransferases (Stanley, 2011). However, in this work we tested the possibility of a direct pass of this set of proteins from the ER avoiding the pass through Golgi, and, concomitantly, it was tried to identify particular proteins susceptible of being targeted through this hypothetical pathway.

The existence of such hypothetical pathway would be significant at biological and biotechnological levels. At a biological level, it would provide a better understanding of protein traffic within cells, and the possible finding of characteristics shared by these proteins would shed light on evolution of targeting processes. Additionally, this information would help finding new models for potential biotechnological applications, as production of recombinant proteins lacking immunogenic α(1,3)fucose and β(1,2)xylose residues without the need of engineered transgenic lines knockout in enzymes involved in plant N-glycan modification.

For high mannose type N-glycoprotein purification, a method based in the affinity of such glycans to the lectin Concanavaline A (Con A) was used. Affinity is really high for high mannose glycans, and its use is widely accepted for specific high mannose glycoprotein purification.

The suitability of the *tgd* mutants for this study relies on their already described alteration of ER-chloroplasts direct transport of molecules (see Chapter 3). Noticeable differences in high mannose glycoprotein levels were found in stroma fractions of *tgd4* mutant (Figs. 3.7 and 3.8), when compared to the wild type and the mutant knocked out in the lipid traslocon complex *tgd1*. In some studies (Andersson *et al.*, 2007; Krause *et al.*, 2012), the existence of *tight membrane contact sites* is considered, although until nowadays it remains unclear whether transport proteins occur in these contact sites. In this line, studies in *complex* plastids have reported that the outermost membrane out of their 4, (i.e., those found in haptophytes, cryptomonads and heterokonts) is continuous with ER and nuclear envelope, and suitable for pass of proteins (Sheiner and Striepen, 2013). In higher plants, there are some works reporting evidences of biochemical continuity of ER and chloroplasts membranes. While some studies suggest the possibility of proteins passing through them (Tan *et al.*, 2011; in *Brassica napus*, and Griffing, 2011 in *Nicotiana tabacum*), others are ambiguous when considering it (Mehrshahi *et al.*, 2013). On the other hand, a protein targeting pathway independent from Golgi, in *Plasmodium* apicoplast has been reported (Tonkin *et al.*, 2006). All the previous data would support the possibility of a direct ER to chloroplast transport. The involvement of these direct ER-chloroplast contact sites

(PLAMs) in *Arabidopsis* protein transport was studied with the *tgd4* mutant, being TGD4 an ER resident protein that seems to be involved in the establishment of the named contact sites between ER and chloroplasts. However, in the experiments performed in this study no-differential protein patterns of *tgd1* compared to the wild type were observed, suggesting that high mannose glycoproteins enter the chloroplast through another traslocon other than TGD permease.

Nevertheless, high mannose glycoproteins could be transported directly from the ER to the chloroplast, through a different protein transport complex still uncharacterized and independent of the TGD permease (*tgd1, tgd2, tgd3*).

In performed assays, no striking differences were detected at cellular level, when comparing wild type and *tgd1* and *tgd4* mutants in immunoblot assays (Figs.3.3 and 3.6). However, appreciable differences were observed in affinoblots of *tgd4* stroma samples when comparing the protein pattern to the wild type. Curiously, those observed differences could not been appreciated using 2-D DIGE (DIGE) followed by MALDI-TOF-MS. DIGE only showed differential expression of canonical chloroplast proteins (with a chloproplast transit peptide) not described as N-glycosylated (Suppl. Table S3.1 and Dataset1 in CD), and it was not possible to identify the differential bands that were observed in affinoblots. This might be due to pleiotropic and side effects of *tgd4* mutation, which alters chloroplasts lipid composition and therefore other internal processes might be affected as well.

Indeed, no significant differences were detected in N-glycosylated proteins at total chloroplast levels using different techniques and approaches, like removal of major chloroplasts proteins as Rubisco (Table 3.1, Suppl.Fig. S3.4), normalization of protein amounts as using the ProteoMiner (Suppl. Table 3.2, in Dataset3 in CD), etc.

In general, it is widely recognized that working in the N-glycosylation area is challenging, due to the difficulties in characterizing the glycans-containing peptides (Shental-Bechor and Levy, 2008). In this work, in addition, the extremely low levels of this type of proteins in the chloroplast, presumably close to the technological detection limit, demonstrates the difficulty of finding them, something also recognized by Zybalov *et al.*, 2008; Zielinska *et al.*, 2012 or Ruiz-May *et al.*, 2012. Therefore more sensitive techniques, not only for the detection of those proteins, but for purification of cell fractions or organelles should be used. Indeed and intriguingly, CAH1 has not been detected by MALDI-TOF/MS (nor further LC-MS/MS) experiments. Although there are strong evidence of CAH1 occurring in the chloroplast (Villarejo *et al.* 2005; Burén *et al.* 2011), MS techniques are unable to detect it, indicating its low abundance, specific expression time and/or the difficulty of identification of certain types of proteins like N-glycosylated ones. Identification failure is probably a combination of sensitivity threshold of the methods used together with low abundance of the protein of interest. The fact that all this study is dealing with the search of very low abundant proteins justify all the techniques employed for purification, isolation or separation and identification of proteins of interest (like Con A affinity columns,

Seppro columns, ProteoMiner, different mass spectrometry approaches…). In order to improve the detection of N-glycoroteins in the chloroplast and achieve some of the points raised above, *Arabidopsis* chloroplast fractions were analyzed by LC-MS/MS in collaboration with a research group specialized in glycoprotein identification. The low amount of significant differential identifications when comparing wild type and *tgd4* in both in chloroplasts and stroma samples, indicates once more that we are looking for proteins which are present in very low amounts, confirming the previously mentioned difficulty of working with N-glycoproteins (Zielinska *et al.* 2012).

Nevertheless, a set of potential chloroplast N-glycoproteins was identified, only after Con A purification and enrichment of the sample, confirmed by LC- MS/MS experiments in presence of labeled water ($H_2{}^{18}O$). The use of labeled water increases the confidence of sites of N-glycosylations as it provokes a higher change of mass of the peptide digested by PNGase (see section 2.6.2 of Material and Methods), facilitating detection by Mass Spectrometer. The proteins that are listed in Table 3.2, are predicted to posses at least one peptide N-glycosylated, therefore being susceptible to be targeted to the chloroplast directly from the endomembrane system. However, except myrosinases TGG1 and TGG2, the rest of proteins are predicted to have transit peptide, which would indicate they are targeted to the chloroplast by the Toc/Tic canonical pathway. Nervertheless, a recent patent by French researchers has described the possibility of chloroplast import of recombinant N-glycoproteins carrying high mannose type glycans in algae by a bipartite targeting sequence (Carlier *et al.* 2014). In this line, and according to Peschke and coworkers (2013), glycosylation residues in combination with a TP may serve as targeting signal. So it could happen that these proteins both are N-glycosylated and possess TP. Therefore the chloroplast proteins identified in Table 3.2 should be subjected to a deeper study focused in how they are transported into this organelle.

It can be set with a high degree of confidence that only a small amount of proteins were described to be located mainly in any part of the cell other than the chloroplast, as dataset provided by Mass Spectrometer was analyzed under three different recognized algorithms: OMSSA (Geer *et al.*, 2004), X!Tandem (Craig and Beavis, 2004), and SEQUEST (Eng *et al.*, 1994), with accurate parameters focused in detection of peptides susceptible of being N-glycosylated (see section 2.6.2 of Material and Methods). Some of the identifications were further confirmed by performing the experiments in presence of $H_2{}^{18}O$, which increases the confidence of identifications, reducing the occurence of false positive (see section 3.4.3 of Chapter 3).

Among all analyzed samples, identification of myrosinaseses TGG1 and TGG2 occurred repeatedly. In the present work, they were identified in independent analyses of individual and highly pure biological replicates. All samples presented low levels of contamination with other subcellular fractions, as the vast majority of proteins identified are chloroplastic (Fig. 3.12). TGGs systematic detection is quite persistent to be considered as a contamination. Presence of

these proteins in *Arabidopsis* chloroplast is a very interesting result. Similarly to CAH1, these proteins contain a more extended C terminus than other proteins of the same family, suggesting that it might contain relevant information. Additionally, they possess an ER targeting signal peptide (SP), although they have been described to be primary resident of stomata cells in vacuoles (Barth and Jander, 2006; Agee *et al.*, 2010). Indeed, to our knowledge, the only analysis using GFP fusions of both, TGG1 and TGG2 located them in the vacuole (Agee *et al.*, 2010). However, several different studies, including the present one, have found them in chloroplasts (Kleffman *et al.*, 2004; Giacomelli *et al.*, 2006; Zybalov *et al.*, 2008; Simm *et al.*, 2013). Therefore, it could be deduced that several different isoforms of these proteins exist within the cell, and differential or dual targeting might be the mechanism used to put them in the proper final destination. In this line, it could be that a small set of proteins are targeted to the chloroplasts, additionally to their main subcellular location, as it has been described for other proteins (Carrie and Small, 2013).

There are several examples in literature in which a protein, which was assumed to belong to one specific cell compartment, has been additionally located or predicted in others, as confirmed by search in any on- line database like SUBA, TAIR, PBD, etc (Carrie *et al.*, 2009; Carrie and Small 2013). Indeed, around 100 proteins have been found in both chloroplast and mitochondrion in plants (Carrie and Small, 2013). A small percentage of plastid proteins are also located elsewhere in the cell, and shared locations with mitochondria, nucleus, peroxisome or cytosol have been described (van Wijk and Baginsky, 2011). For example, the study performed by Krause *et al.*, in 2012 provides evidence of proteins being targeted to both chloroplast and nucleus, suggesting they are part of cell signaling system as response to external variations. Maybe this potential dual targeting occurs only under certain conditions, still to be defined. For example, there are several evidences of plastids being involved in recognition of pathogens and responses to wounding. As TGGs are known to be involved in defense (Barth and Jander, 2006), maybe dual targeting could be a cellular response to changing environment.

This dual targeting and specific cell response to different conditions is in agreement with the increasing acceptance of dynamism of membranes (see section 4 of Introduction), which could assist punctual proteome enrichment, or divert a set of proteins to other cell compartments depending on the needs. Another option could be that those identified myrosinases are alternative isoforms still not described. In addition to TGGs, the proteins listed in Table 3.2 were also identified under the same experimental conditions. Although several studies have explored chloroplast proteome, to our knowledge, this is the first time that has been performed a specific proteomic analysis of *Arabidopsis* chloroplast fractions, focused in the identification of HMGP susceptible of being transported directly from ER, by comparison with a mutant, such as *tgd4*, which have disrupted a direct ER-chloroplast import route of molecules. Additionally, this is the first time that the presence of high mannose glycan proteins in *Arabidopsis* chloroplasts has been

reported by different approaches. Furthermore, N-glycoproteome study including $H_2{}^{18}O$ labeling permits higher reliability in N-glycosylation sites identification (Küster and Mann, 1999; Zielinska *et al.*, 2010; Mathieu -Rivet *et al.*, 2013). However, deeper studies on how these proteins reach the chloroplast should be performed. For example, describing the exact function of *tgd4* protein in PLAMs formation, which targeting mechanisms these proteins use, or how they are transported across the organelle membranes and elements involved remain to be described.

On one side, it would confirm the existence of chloroplast protein import pathways other than canonical one (for example, through and undescribed translocon formed in the PLAMs), complementing previous studies on N-glycoproteins in plastid of higher plants. On the biotechnological area, confirmation that at least a set of high mannose N-glycoproteins are targeted to the chloroplasts, open a door for new forms of production of non-allergic recombinant proteins.

X. CONCLUSIONS

1. N-glycosylation of CAH1 affects protein folding and trafficking, therefore influencing its function when transiently expressed in *Arabidopsis* protoplasts and *Nicotiana benthamiana* leaves.
2. Formation of a disulphide bond between Cys27 and Cys191 (C1 and C3) of CAH1 was confirmed in *Arabidopsis* protoplasts. Disruption of the bond due to cysteine mutation leads to high mannose N-glycan enrichment, indicating some influence in protein folding and additional retention in the ER.
3. Transient expression in both *Arabidopsis* protoplasts and *Nicotiana benthamiana* leaves showed that mutation of two lysine residues in CAH1 C terminus does not affect N-glycosylation or proper folding of the protein. However, complete deletion of the C terminus alters N-glycosylation status, increasing high mannose content. ER-retention of the protein is not due to defects in protein folding, since the ER quality control system (ERQC) is not induced, suggesting that an alternative unknown mechanism involving the C terminus of the protein is required for ER export.
4. Although lack of statistical significance hampers unequivocal conclusions, general trends show that N-glycosylation defects negatively influence plant root development and and cell wall integrity under salt stress.This effect seems more pronounced when complete complex N-glycans are missing, compared to plants lacking only α(1,3)fucose residues. These moderate trends would require further confirmation. Similarly, neither the type of N-glycan attached to KOR1, nor the increasing stress produced by salt treatments, seem to directly affect gene expression level.
5. Subcellular localization of *Arabidopsis thaliana* α(1,3)fucosyltransferases (FT11 and FT12) was not possible with the immunohistochemistry and microscopy approaches used in this work. Different techniques should be used in order to determine the exact subcellular localization of these enzymes.
6. The existence of high mannose N-glycoproteins in *Arabidopsis thaliana* chloroplast has been confirmed for the first time, opening new opportunities for possible future biotechnological applications. Nevertheless, occurrence of N-glycoproteins in general, and high mannose N-glycoproteins in particular, is extremely minor in *Arabidopsis* chloroplast.
7. High mannose N-glycoprotein transport does not depend on the lipid permease complex TGD located in the inner envelope membrane of the chloroplast, since N-glycoprotein import is not affected in the *tgd1* mutant. However, further analyses should be performed in order to determine TGD4 implication in the mechanisms that regulate N-glycoprotein targeting and trafficking.

8. Myrosinases TGG1 and TGG2 were found in chloroplasts preparations and among all replicates analyzed by different experimental approaches. Therefore they could be considered as potential models for plastid N-glycoprotein analyses. Nevertheless, additional studies should be performed in order to further corroborate this fact and its biological implications or significance.

X. CONCLUSIONES

1. La N- glicosilación de CAH1 afecta al plegamiento y tráfico de la proteína, influyendo por tanto en su función, cuando se expresa de forma transitoria en protoplastos de *Arabidopsis* y en hojas de *Nicotiana benthamiana*.
2. La formación de un enlace disulfuro entre las Cys27 y Cys191 (C1 y C3) de CAH1 ha sido confirmada en protoplastos de *Arabidopsis*. La rotura del puente disulfuro a causa de las mutaciones en estas cisteínas, conduce al enriquecimiento en N-glicanos ricos en manosa en las proteínas, lo que indicaría su influencia en el plegamiento de proteínas y su consecuente retención en el retículo endoplasmático.
3. La expresión transitoria en protoplastos de *Arabidopsis* y hojas de *Nicotiana benthamiana* de construcciones de CAH1 con alteraciones en su C terminal, demostró, por un lado, que la mutación de dos residuos de lisina en el C terminal de CAH1no afecta a su N-glicosilación o correcto plegamiento. Por otro lado, sin embargo, la eliminación completa del C terminal altera el estado de N-glicosilación, aumentando el contenido de glicanos ricos en manosa. La retención de la proteína en el retículo no se debe a defectos en el plegamiento de la proteína, ya que el sistema de control de calidad del retículo (ERAD) no está inducido en estas células. Esto sugiere que el extremo C terminal de la proteína sería necesario para la exportación, a través de un mecanismo todavía desconocido, desde el ER.
4. Aunque la falta de significación estadística dificulta la obtención de conclusiones determinantes, los resultados parecen indicar que los defectos en N-glicosilación influyen negativamente en el desarrollo radicular y en la integridad de las paredes celulares de plantas bajo estrés salino. Este efecto parece ser más pronunciado en el caso de que la carencia de residuos de N-glicano complejos sea completa, en comparación con plantas que sólo carecen de residuos de α(1,3)fucosa. Sin embargo, es necesario confirmar dichas tendencias. Por otro lado, la expresión del gen que codifica la N-glicoproteína implicada en la formación de pared celular KOR1, no parece verse directamente alterada ni por el estrés salino al que se sometió a las plantas, ni por el tipo de N- glicano unido a la estructura de la proteína.
5. La localización subcelular precisa de las α(1,3)fucosiltransferasas (FT11 y FT12) de *Arabidopsis thaliana* no fue posible con los estudios de inmunohistoquímica y microscopía utilizados en este trabajo. Sería necesario un nuevo planteamiento que implique utilizar técnicas distintas, con el fin de determinar la localización subcelular exacta de estas enzimas.
6. Se ha confirmado por primera vez la existencia de N-glicoproteínas ricas en manosa en el cloroplasto de *Arabidopsis thaliana*, abriendo nuevas oportunidades para posibles aplicaciones biotecnológicas en el futuro. Sin embargo,la presencia de N- glicoproteínas en

general, y particularmente N- glicoproteínas ricas en manosa, es extremadamente minoritaria en el cloroplasto de *Arabidopsis*.

7. El transporte de N-glicoproteínas ricas en manosa no depende de la permeasa TGD situada en la membrana interna de los cloroplastos, e implicada en la importación de precursores lipídicos, ya que la entrada de N–glicoproteínas a dicho orgánulo no se ve afectada en el mutante *tgd1*. Sin embargo, determinar el papel desempeñado por la proteína TGD4 en los mecanismos que regulan el tráfico y localización de N-glicoproteínas requiere realizar análisis adicionales.
8. Las mirosinasas TGG1 y TGG2 se encontraron en preparaciones de cloroplastos y en todas las réplicas analizadas mediante distintas técnicas experimentales. Por lo tanto, podrían ser consideradas como posibles modelos para el análisis N-glicoproteínas cloroplásticas. Sin embargo, aún deben realizarse más estudios para corroborar este hecho y sus posibles implicaciones biológicas.

XI. REFERENCES

Achleitner, G., B. Gaigg, A. Krasser, E. Kainersdorfer, S. D. Kohlwein, A. Perktold, G. Zellnig and G. Daum (1999) Association between the endoplasmic reticulum and mitochondria of yeast facilitates interorganelle transport of phospholipids through membrane contact. *European journal of biochemistry / FEBS*. **264**, 545-53.

Agee, A. E., M. Surpin, E. J. Sohn, T. Girke, A. Rosado, B. W. Kram, C. Carter, A. M. Wentzell, D. J. Kliebenstein, H. C. Jin, O. K. Park, H. Jin, G. R. Hicks and N. V. Raikhel (2010) MODIFIED VACUOLE PHENOTYPE1 is an Arabidopsis myrosinase-associated protein involved in endomembrane protein trafficking. *Plant physiology*. **152**, 120-32.

Andersson, M. X., M. Goksor and A. S. Sandelius (2007) Optical manipulation reveals strong attracting forces at membrane contact sites between endoplasmic reticulum and chloroplasts. *The Journal of biological chemistry*. **282**, 1170-4.

Armbruster, U., A. Hertle, E. Makarenko, J. Zuhlke, M. Pribil, A. Dietzmann, I. Schliebner, E. Aseeva, E. Fenino, M. Scharfenberg, C. Voigt and D. Leister (2009) Chloroplast proteins without cleavable transit peptides: rare exceptions or a major constituent of the chloroplast proteome? *Molecular plant*. **2**, 1325-35.

Asatsuma, S., C. Sawada, K. Itoh, M. Okito, A. Kitajima and T. Mitsui (2005) Involvement of alpha-amylase I-1 in starch degradation in rice chloroplasts. *Plant & cell physiology*. **46**, 858-69.

Awai, K., C. Xu, B. Tamot and C. Benning (2006) A phosphatidic acid-binding protein of the chloroplast inner envelope membrane involved in lipid trafficking. *Proceedings of the National Academy of Sciences of the United States of America*. **103**, 10817-22.

Bakker, H., E. Schijlen, T. de Vries, W. E. Schiphorst, W. Jordi, A. Lommen, D. Bosch and I. van Die (2001) Plant members of the alpha1-->3/4-fucosyltransferase gene family encode an alpha1-->4-fucosyltransferase, potentially involved in Lewis(a) biosynthesis, and two core alpha1-->3-fucosyltransferases. *FEBS letters*. **507**, 307-12.

Bardor, M., G. Cabrera, J. Stadlmann, P. Lerouge, J. A. Cremata, V. Gomord and A. C. Fitchette (2009) N-glycosylation of plant recombinant pharmaceuticals. *Methods in molecular biology*. **483**, 239-64.

Bargmann, B. O. and K. D. Birnbaum (2010) Fluorescence activated cell sorting of plant protoplasts. *J Vis Exp*.

Barth, C. and G. Jander (2006) Arabidopsis myrosinases TGG1 and TGG2 have redundant function in glucosinolate breakdown and insect defense. *The Plant journal : for cell and molecular biology*. **46**, 549-62.

Bates, S., H. B. Hughes, C. A. Munro, W. P. Thomas, D. M. MacCallum, G. Bertram, A. Atrih, M. A. Ferguson, A. J. Brown, F. C. Odds and N. A. Gow (2006) Outer chain N-glycans are required for cell wall integrity and virulence of Candida albicans. *The Journal of biological chemistry*. **281**, 90-8.

Ben-Menachem, R., N. Regev-Rudzki and O. Pines (2011) The aconitase C-terminal domain is an independent dual targeting element. *Journal of molecular biology*. **409**, 113-23.

Bencur, P., H. Steinkellner, B. Svoboda, J. Mucha, R. Strasser, D. Kolarich, S. Hann, G. Kollensperger, J. Glossl, F. Altmann and L. Mach (2005) Arabidopsis thaliana beta1,2-xylosyltransferase: an unusual glycosyltransferase with the potential to act at multiple stages of the plant N-glycosylation pathway. *The Biochemical journal*. **388**, 515-25.

Bencurova, M., W. Hemmer, M. Focke-Tejkl, I. B. Wilson and F. Altmann (2004) Specificity of IgG and IgE antibodies against plant and insect glycoprotein glycans determined with artificial glycoforms of human transferrin. *Glycobiology*. **14**, 457-66.

Benning, C. (2009) Mechanisms of lipid transport involved in organelle biogenesis in plant cells. *Annual review of cell and developmental biology*. **25**, 71-91.

Benning, C., C. Xu and K. Awai (2006) Non-vesicular and vesicular lipid trafficking involving plastids. *Current opinion in plant biology*. **9**, 241-7.

Betz, S. F. (1993) Disulfide bonds and the stability of globular proteins. *Protein science : a publication of the Protein Society*. **2**, 1551-8.

Bhattacharya, D., J. M. Archibald, A. P. Weber and A. Reyes-Prieto (2007) How do endosymbionts become organelles? Understanding early events in plastid evolution. *BioEssays : news and reviews in molecular, cellular and developmental biology*. **29**, 1239-46.

Blum, H., Beier,H. and Gross,H.J. Improved silver staining of plant proteins, RNA and DNA in polyacrylamide gels. Electrophoresis (1987) 8, 93-99.

Boisson, M., V. Gomord, C. Audran, N. Berger, B. Dubreucq, F. Granier, P. Lerouge, L. Faye, M. Caboche and L. Lepiniec (2001) Arabidopsis glucosidase I mutants reveal a critical role of N-glycan trimming in seed development. *The EMBO journal*. **20**, 1010-9.

Bosch, D., A. Castilho, A. Loos, A. Schots and H. Steinkellner (2013) N-glycosylation of plant-produced recombinant proteins. *Current pharmaceutical design*. **19**, 5503-12.

Both, P. (2010) Etude structure-fuction d'une fucosyltransfersase (FucT-A) de *Arabidopsis thaliana*, Univertié Joseph Fourier (Grenoble I)

Both, P., L. Sobczak, C. Breton, S. Hann, K. Nobauer, K. Paschinger, S. Kozmon, J. Mucha and I. B. Wilson (2011) Distantly related plant and nematode core alpha1,3-fucosyltransferases display similar trends in structure-function relationships. *Glycobiology*. **21**, 1401-15.

Boyle, J., Lehninger principles of biochemistry (4th ed.): Nelson, D., and Cox, M. Biochem. Mol. Biol. Educ., 33: 74–75. doi: 10.1002/bmb.2005.494033010419

Bradford, M. M. (1976) A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Analytical biochemistry*. **72**, 248-54.

Buren, S., C. Ortega-Villasante, A. Blanco-Rivero, A. Martinez-Bernardini, T. Shutova, D. Shevela, J. Messinger, L. Bako, A. Villarejo and G. Samuelsson (2011) Importance of post-translational modifications for functionality of a chloroplast-localized carbonic anhydrase (CAH1) in Arabidopsis thaliana. *PLoS One*. **6**, e21021.

Burn, J. E., U. A. Hurley, R. J. Birch, T. Arioli, A. Cork and R. E. Williamson (2002) The cellulose-deficient Arabidopsis mutant rsw3 is defective in a gene encoding a putative glucosidase II, an enzyme processing N-glycans during ER quality control. *The Plant journal : for cell and molecular biology*. **32**, 949-60.

Carlier, A., R. Michel, J. P. Cadoret, A. Lejeune and N. Dufourmantel (2014) Production of high mannose glycosylated proteins stored in the plastid of microalgae, Google Patents

Carrie, C., K. Kuhn, M. W. Murcha, O. Duncan, I. D. Small, N. O'Toole and J. Whelan (2009) Approaches to defining dual-targeted proteins in Arabidopsis. *The Plant journal : for cell and molecular biology*. **57**, 1128-39.

Carrie, C. and I. Small (2013) A reevaluation of dual-targeting of proteins to mitochondria and chloroplasts. *Biochim Biophys Acta*. **1833**, 253-9.

Castilho, A. and H. Steinkellner (2012) Glyco-engineering in plants to produce human-like N-glycan structures. *Biotechnology journal*. **7**, 1088-98.

Cavalier-Smith, T. (2009) Predation and eukaryote cell origins: a coevolutionary perspective. *Int J Biochem Cell Biol*. **41**, 307-22.

Coutinho, P. M., E. Deleury, G. J. Davies and B. Henrissat (2003) An Evolving Hierarchical Family Classification for Glycosyltransferases. *Journal of molecular biology*. **328**, 307-317.

Craciun, A. R., M. Courbot, F. Bourgis, P. Salis, P. Saumitou-Laprade and N. Verbruggen (2006) Comparative cDNA-AFLP analysis of Cd-tolerant and -sensitive genotypes derived from crosses between the Cd hyperaccumulator Arabidopsis halleri and Arabidopsis lyrata ssp. petraea. *Journal of experimental botany*. **57**, 2967-83.

Craig, R. and R. C. Beavis (2004) TANDEM: matching proteins with tandem mass spectra. *Bioinformatics*. **20**, 1466-7.

Cubero, B., Y. Nakagawa, X. Y. Jiang, K. J. Miura, F. Li, K. G. Raghothama, R. A. Bressan, P. M. Hasegawa and J. M. Pardo (2009) The phosphate transporter PHT4;6 is a determinant of salt tolerance that is localized to the Golgi apparatus of Arabidopsis. *Molecular plant*. **2**, 535-52.

Curtis, M. D. and U. Grossniklaus (2003) A gateway cloning vector set for high-throughput functional analysis of genes in planta. *Plant physiology*. **133**, 462-9.

Chen, J. K., J. Taipale, M. K. Cooper and P. A. Beachy (2002) Inhibition of Hedgehog signaling by direct binding of cyclopamine to Smoothened. *Genes & development*. **16**, 2743-8.

Chen, M. H., L. F. Huang, H. M. Li, Y. R. Chen and S. M. Yu (2004) Signal peptide-dependent targeting of a rice alpha-amylase and cargo proteins to plastids and extracellular compartments of plant cells. *Plant physiology*. **135**, 1367-77.

Daniell, H., S. Chebolu, S. Kumar, M. Singleton and R. Falconer (2005) Chloroplast-derived vaccine antigens and other therapeutic proteins. *Vaccine*. **23**, 1779-83.

Daniell, H., N. D. Singh, H. Mason and S. J. Streatfield (2009) Plant-made vaccine antigens and biopharmaceuticals. *Trends Plant Sci*. **14**, 669-79.

Davoodi-Semiromi, A., N. Samson and H. Daniell (2009) The green vaccine: A global strategy to combat infectious and autoimmune diseases. *Human vaccines*. **5**, 488-93.

Davoodi-Semiromi, A., M. Schreiber, S. Nalapalli, D. Verma, N. D. Singh, R. K. Banks, D. Chakrabarti and H. Daniell (2010) Chloroplast-derived vaccine antigens confer dual immunity against cholera and malaria by oral or injectable delivery. *Plant biotechnology journal*. **8**, 223-42.

Douzery, E. J., E. A. Snell, E. Bapteste, F. Delsuc and H. Philippe (2004) The timing of eukaryotic evolution: does a relaxed molecular clock reconcile proteins and fossils? *Proceedings of the National Academy of Sciences of the United States of America*. **101**, 15386-91.

Durocher, Y. and M. Butler (2009) Expression systems for therapeutic glycoprotein production. *Curr Opin Biotechnol*. **20**, 700-7.

Endler, A. and S. Persson (2011) Cellulose synthases and synthesis in Arabidopsis. *Molecular plant*. **4**, 199-211.

Etzler, M. E. and D. Mohnen (2009) Viridiplantae. *Essentials of Glycobiology*. Varki, A., Cummings, R. D., Esko, J. D.et al (eds.) Cold Spring Harbor (NY)

Fabre, N., I. M. Reiter, N. Becuwe-Linka, B. Genty and D. Rumeau (2007) Characterization and expression analysis of genes encoding alpha and beta carbonic anhydrases in Arabidopsis. *Plant Cell Environ*. **30**, 617-29.

Faveeuw, C., T. Mallevaey, K. Paschinger, I. B. Wilson, J. Fontaine, R. Mollicone, R. Oriol, F. Altmann, P. Lerouge, M. Capron and F. Trottein (2003) Schistosome N-glycans containing core alpha 3-fucose and core beta 2-xylose epitopes are strong inducers of Th2 responses in mice. *Eur J Immunol*. **33**, 1271-81.

Faye, L. and M. J. Chrispeels (1985) Characterization of N-linked oligosaccharides by affinoblotting with concanavalin A-peroxidase and treatment of the blots with glycosidases. *Analytical biochemistry*. **149**, 218-24.

Fernandez-San Millan, A., S. M. Ortigosa, S. Hervas-Stubbs, P. Corral-Martinez, J. M. Segui-Simarro, J. Gaetan, P. Coursaget and J. Veramendi (2008) Human papillomavirus L1 protein expressed in tobacco chloroplasts self-assembles into virus-like particles that are highly immunogenic. *Plant biotechnology journal*. **6**, 427-41.

Foetisch, K., S. Westphal, I. Lauer, M. Retzek, F. Altmann, D. Kolarich, S. Scheurer and S. Vieths (2003) Biological activity of IgE specific for cross-reactive carbohydrate determinants. *The Journal of allergy and clinical immunology*. **111**, 889-96.

Forsayeth, J. R., Y. Gu and Z. W. Hall (1992) BiP forms stable complexes with unassembled subunits of the acetylcholine receptor in transfected COS cells and in C2 muscle cells. *The Journal of cell biology*. **117**, 841-7.

Foth, B. J., S. A. Ralph, C. J. Tonkin, N. S. Struck, M. Fraunholz, D. S. Roos, A. F. Cowman and G. I. McFadden (2003) Dissecting apicoplast targeting in the malaria parasite Plasmodium falciparum. *Science*. **299**, 705-8.

Freeze, H. H. and M. Aebi (2005) Altered glycan structures: the molecular basis of congenital disorders of glycosylation. *Curr Opin Struct Biol*. **15**, 490-8.

Friso, G., L. Giacomelli, A. J. Ytterberg, J. B. Peltier, A. Rudella, Q. Sun and K. J. Wijk (2004) In-depth analysis of the thylakoid membrane proteome of Arabidopsis thaliana chloroplasts: new proteins, new functions, and a plastid proteome database. *The Plant cell*. **16**, 478-99.

Gagat, P., A. Bodyl and P. Mackiewicz (2013) How protein targeting to primary plastids via the endomembrane system could have evolved? A new hypothesis based on phylogenetic studies. *Biol Direct*. **8**, 18.

Gaikwad, A., K. K. Tewari, D. Kumar, W. Chen and S. K. Mukherjee (1999) Isolation and characterisation of the cDNA encoding a glycosylated accessory protein of pea chloroplast DNA polymerase. *Nucleic acids research*. **27**, 3120-9.

Gao, M., Y. Li, X. Xue, X. Wang and J. Long (2012) Stable plastid transformation for high-level recombinant protein expression: promises and challenges. *J Biomed Biotechnol*. **2012**, 158232.

Giacomelli, L., A. Rudella and K. J. van Wijk (2006) High light response of the thylakoid proteome in arabidopsis wild type and the ascorbate-deficient mutant vtc2-2. A comparative proteomics study. *Plant physiology*. **141**, 685-701.

Gillmor, C. S., P. Poindexter, J. Lorieau, M. M. Palcic and C. Somerville (2002) Alpha-glucosidase I is required for cellulose biosynthesis and morphogenesis in Arabidopsis. *The Journal of cell biology*. **156**, 1003-13.

Gomord, V., A. C. Fitchette, L. Menu-Bouaouiche, C. Saint-Jore-Dupas, C. Plasson, D. Michaud and L. Faye (2010) Plant-specific glycosylation patterns in the context of therapeutic protein production. *Plant biotechnology journal*. **8**, 564-87.

Goodin, M. M., D. Zaitlin, R. A. Naidu and S. A. Lommel (2008) Nicotiana benthamiana: its history and future as a model for plant-pathogen interactions. *Molecular plant-microbe interactions : MPMI*. **21**, 1015-26.

Goulas, E., M. Schubert, T. Kieselbach, L. A. Kleczkowski, P. Gardestrom, W. Schroder and V. Hurry (2006) The chloroplast lumen and stromal proteomes of Arabidopsis thaliana show differential sensitivity to short- and long-term exposure to low temperature. *The Plant journal : for cell and molecular biology*. **47**, 720-34.

Gould, S. B., R. F. Waller and G. I. McFadden (2008) Plastid evolution. *Annual review of plant biology*. **59**, 491-517.

Griffing, L. R. (2011) Laser stimulation of the chloroplast/endoplasmic reticulum nexus in tobacco transiently produces protein aggregates (boluses) within the endoplasmic reticulum and stimulates local ER remodeling. *Molecular plant*. **4**, 886-95.

Guex, N. and M. C. Peitsch (1997) SWISS-MODEL and the Swiss-PdbViewer: an environment for comparative protein modeling. *Electrophoresis*. **18**, 2714-23.

Han, H. M., C. Bouchet-Marquis, J. Huebinger and M. Grabenbauer (2013) Golgi apparatus analyzed by cryo-electron microscopy. *Histochem Cell Biol*. **140**, 369-81.

Hawes, C. (2012) The ER/Golgi Interface - Is There Anything in-between? *Front Plant Sci*. **3**, 73.

Hayashi, M., A. Tsuru, T. Mitsui, N. Takahashi, H. Hanzawa, Y. Arata and T. Akazawa (1990) Structure and biosynthesis of the xylose-containing carbohydrate moiety of rice alpha-amylase. *European journal of biochemistry / FEBS*. **191**, 287-95.

Hebert, D. N. and M. Molinari (2007) In and out of the ER: protein folding, quality control, degradation, and related human diseases. *Physiological reviews*. **87**, 1377-408.

Hilvo, M., M. Tolvanen, A. Clark, B. Shen, G. N. Shah, A. Waheed, P. Halmi, M. Hanninen, J. M. Hamalainen, M. Vihinen, W. S. Sly and S. Parkkila (2005) Characterization of CA XV, a new GPI-anchored form of carbonic anhydrase. *The Biochemical journal*. **392**, 83-92.

His, I., A. Driouich, F. Nicol, A. Jauneau and H. Hofte (2001) Altered pectin composition in primary cell walls of korrigan, a dwarf mutant of Arabidopsis deficient in a membrane-bound endo-1,4-beta-glucanase. *Planta*. **212**, 348-58.

Hong, Z., H. Jin, T. Tzfira and J. Li (2008) Multiple mechanism-mediated retention of a defective brassinosteroid receptor in the endoplasmic reticulum of Arabidopsis. *The Plant cell*. **20**, 3418-29.

Hossain, Z., M. Z. Nouri and S. Komatsu (2012) Plant cell organelle proteomics in response to abiotic stress. *Journal of proteome research*. **11**, 37-48.

Huber, L. A., K. Pfaller and I. Vietor (2003) Organelle proteomics: implications for subcellular fractionation in proteomics. *Circ Res*. **92**, 962-8.

Ioffe, E., Y. Liu and P. Stanley (1996) Essential role for complex N-glycans in forming an organized layer of bronchial epithelium. *Proceedings of the National Academy of Sciences of the United States of America*. **93**, 11041-6.

Ioffe, E. and P. Stanley (1994) Mice lacking N-acetylglucosaminyltransferase I activity die at mid-gestation, revealing an essential role for complex or hybrid N-linked carbohydrates. *Proceedings of the National Academy of Sciences of the United States of America*. **91**, 728-32.

Jez, J., A. Castilho, J. Grass, K. Vorauer-Uhl, T. Sterovsky, F. Altmann and H. Steinkellner (2013) Expression of functionally active sialylated human erythropoietin in plants. *Biotechnology journal*. **8**, 371-82.

Kajiura, H., T. Okamoto, R. Misaki, Y. Matsuura and K. Fujiyama (2012) Arabidopsis beta1,2-xylosyltransferase: substrate specificity and participation in the plant-specific N-glycosylation pathway. *J Biosci Bioeng*. **113**, 48-54.

Kang, J. S., J. Frank, C. H. Kang, H. Kajiura, M. Vikram, A. Ueda, S. Kim, J. D. Bahk, B. Triplett, K. Fujiyama, S. Y. Lee, A. von Schaewen and H. Koiwa (2008) Salt tolerance of Arabidopsis thaliana requires maturation of N-glycosylated proteins in the Golgi apparatus. *Proceedings of the National Academy of Sciences of the United States of America*. **105**, 5933-8.

Karniely, S. and O. Pines (2005) Single translation--dual destination: mechanisms of dual protein targeting in eukaryotes. *EMBO Rep*. **6**, 420-5.

Kaulfurst-Soboll, H., S. Rips, H. Koiwa, H. Kajiura, K. Fujiyama and A. von Schaewen (2011) Reduced immunogenicity of Arabidopsis hgl1 mutant N-glycans caused by altered accessibility of xylose and core fucose epitopes. *The Journal of biological chemistry*. **286**, 22955-64.

Kaushik, S., D. Mohanty and A. Surolia (2011) Role of glycosylation in structure and stability of Erythrina corallodendron lectin (EcorL): a molecular dynamics study. *Protein science : a publication of the Protein Society*. **20**, 465-81.

Kelleher, D. J. and R. Gilmore (2006) An evolving view of the eukaryotic oligosaccharyltransferase. *Glycobiology*. **16**, 47R-62R.

Kelly, A. A. and P. Dormann (2004) Green light for galactolipid trafficking. *Current opinion in plant biology*. **7**, 262-9.

Kiefer, F., K. Arnold, M. Kunzli, L. Bordoli and T. Schwede (2009) The SWISS-MODEL Repository and associated resources. *Nucleic acids research*. **37**, D387-92.

Kim, Y. C., N. Jahren, M. D. Stone, N. D. Udeshi, T. W. Markowski, B. A. Witthuhn, J. Shabanowitz, D. F. Hunt and N. E. Olszewski (2013) Identification and origin of N-linked beta-D-N-acetylglucosamine monosaccharide modifications on Arabidopsis proteins. *Plant physiology*. **161**, 455-64.

Kitajima, A., S. Asatsuma, H. Okada, Y. Hamada, K. Kaneko, Y. Nanjo, Y. Kawagoe, K. Toyooka, K. Matsuoka, M. Takeuchi, A. Nakano and T. Mitsui (2009) The rice alpha-amylase glycoprotein is targeted from the Golgi apparatus through the secretory pathway to the plastids. *The Plant cell*. **21**, 2844-58.

Kjaer, S. and C. F. Ibanez (2003) Intrinsic susceptibility to misfolding of a hot-spot for Hirschsprung disease mutations in the ectodomain of RET. *Hum Mol Genet*. **12**, 2133-44.

Kleffmann, T., D. Russenberger, A. von Zychlinski, W. Christopher, K. Sjolander, W. Gruissem and S. Baginsky (2004) The Arabidopsis thaliana chloroplast proteome reveals pathway abundance and novel protein functions. *Curr Biol*. **14**, 354-62.

Kohorn, B. D., M. Kobayashi, S. Johansen, H. P. Friedman, A. Fischer and N. Byers (2006) Wall-associated kinase 1 (WAK1) is crosslinked in endomembranes, and transport to the cell surface requires correct cell-wall synthesis. *Journal of cell science*. **119**, 2282-90.

Koiwa, H., F. Li, M. G. McCully, I. Mendoza, N. Koizumi, Y. Manabe, Y. Nakagawa, J. Zhu, A. Rus, J. M. Pardo, R. A. Bressan and P. M. Hasegawa (2003) The STT3a subunit isoform of the Arabidopsis oligosaccharyltransferase controls adaptive responses to salt/osmotic stress. *The Plant cell*. **15**, 2273-84.

Koprivova, A., C. Stemmer, F. Altmann, A. Hoffmann, S. Kopriva, G. Gorr, R. Reski and E. L. Decker (2004) Targeted knockouts of Physcomitrella lacking plant-specific immunogenic N-glycans. *Plant biotechnology journal*. **2**, 517-23.

Krause, K., S. Oetke and K. Krupinska (2012) Dual targeting and retrograde translocation: regulators of plant nuclear gene expression can be sequestered by plastids. *Int J Mol Sci*. **13**, 11085-101.

Kwok, E. Y. and M. R. Hanson (2004) Stromules and the dynamic nature of plastid morphology. *Journal of microscopy*. **214**, 124-37.

Kwon, H. K., R. Yokoyama and K. Nishitani (2005) A proteomic approach to apoplastic proteins involved in cell wall regeneration in protoplasts of Arabidopsis suspension-cultured cells. *Plant & cell physiology*. **46**, 843-57.

Lamesch, P., T. Z. Berardini, D. Li, D. Swarbreck, C. Wilks, R. Sasidharan, R. Muller, K. Dreher, D. L. Alexander, M. Garcia-Hernandez, A. S. Karthikeyan, C. H. Lee, W. D. Nelson, L. Ploetz, S. Singh, A. Wensel and E. Huala (2012) The Arabidopsis Information Resource (TAIR): improved gene annotation and new tools. *Nucleic acids research*. **40**, D1202-10.

Lane, D. R., A. Wiedemeier, L. Peng, H. Hofte, S. Vernhettes, T. Desprez, C. H. Hocart, R. J. Birch, T. I. Baskin, J. E. Burn, T. Arioli, A. S. Betzner and R. E. Williamson (2001) Temperature-sensitive alleles of RSW2 link the KORRIGAN endo-1,4-beta-glucanase to cellulose synthesis and cytokinesis in Arabidopsis. *Plant physiology*. **126**, 278-88.

Larkin, A. and B. Imperiali (2011) The expanding horizons of asparagine-linked glycosylation. *Biochemistry*. **50**, 4411-26.

Lau, K. S., E. A. Partridge, A. Grigorian, C. I. Silvescu, V. N. Reinhold, M. Demetriou and J. W. Dennis (2007) Complex N-glycan number and degree of branching cooperate to regulate cell proliferation and differentiation. *Cell*. **129**, 123-34.

Lee, M. and Y. Yang (2006) Transient Expression Assay by Agroinfiltration of Leaves. *Arabidopsis Protocols*. Salinas, J. and Sanchez-Serrano, J. (eds.), pp 225-229, Humana Press

Lerich, A., S. Hillmer, M. Langhans, D. Scheuring, P. van Bentum and D. G. Robinson (2012) ER Import Sites and Their Relationship to ER Exit Sites: A New Model for Bidirectional ER-Golgi Transport in Higher Plants. *Front Plant Sci*. **3**, 143.

Lerouge, P., M. Cabanes-Macheteau, C. Rayon, A. C. Fischette-Laine, V. Gomord and L. Faye (1998) N-glycoprotein biosynthesis in plants: recent developments and future trends. *Plant molecular biology*. **38**, 31-48.

Lerouxel, O., G. Mouille, C. Andeme-Onzighi, M. P. Bruyant, M. Seveno, C. Loutelier-Bourhis, A. Driouich, H. Hofte and P. Lerouge (2005) Mutants in DEFECTIVE GLYCOSYLATION, an Arabidopsis homolog of an oligosaccharyltransferase complex subunit, show protein

underglycosylation and defects in cell differentiation and growth. *The Plant journal : for cell and molecular biology*. **42**, 455-68.

Levitan, A., T. Trebitsh, V. Kiss, Y. Pereg, I. Dangoor and A. Danon (2005) Dual targeting of the protein disulfide isomerase RB60 to the chloroplast and the endoplasmic reticulum. *Proceedings of the National Academy of Sciences of the United States of America*. **102**, 6225-30.

Li, H. M. and C. C. Chiu (2010) Protein transport into chloroplasts. *Annual review of plant biology*. **61**, 157-80.

Li, J., C. Zhao-Hui, M. Batoux, V. Nekrasov, M. Roux, D. Chinchilla, C. Zipfel and J. D. Jones (2009) Specific ER quality control components required for biogenesis of the plant innate immune receptor EFR. *Proceedings of the National Academy of Sciences of the United States of America*. **106**, 15973-8.

Liebminger, E., J. Grass, F. Altmann, L. Mach and R. Strasser (2013) Characterizing the link between glycosylation state and enzymatic activity of the endo-beta1,4-glucanase KORRIGAN1 from Arabidopsis thaliana. *The Journal of biological chemistry*. **288**, 22270-80.

Liebminger, E., J. Grass, J. Jez, L. Neumann, F. Altmann and R. Strasser (2012) Myrosinases TGG1 and TGG2 from Arabidopsis thaliana contain exclusively oligomannosidic N-glycans. *Phytochemistry*. **84**, 24-30.

Lowe, J. B. and J. D. Marth (2003) A genetic approach to Mammalian glycan function. *Annu Rev Biochem*. **72**, 643-91.

Lukowitz, W., T. C. Nickle, D. W. Meinke, R. L. Last, P. L. Conklin and C. R. Somerville (2001) Arabidopsis cyt1 mutants are deficient in a mannose-1-phosphate guanylyltransferase and point to a requirement of N-linked glycosylation for cellulose biosynthesis. *Proceedings of the National Academy of Sciences of the United States of America*. **98**, 2262-7.

Ma, B., J. L. Simala-Grant and D. E. Taylor (2006) Fucosylation in prokaryotes and eukaryotes. *Glycobiology*. **16**, 158R-184R.

Mallick, P. and B. Kuster (2010) Proteomics: a pragmatic perspective. *Nature biotechnology*. **28**, 695-709.

Mathieu-Rivet, E., M. Scholz, C. Arias, F. Dardelle, S. Schulze, F. Le Mauff, G. Teo, A. K. Hochmal, A. Blanco-Rivero, C. Loutelier-Bourhis, M. C. Kiefer-Meyer, C. Fufezan, C. Burel, P. Lerouge, F. Martinez, M. Bardor and M. Hippler (2013) Exploring the N-glycosylation pathway in Chlamydomonas reinhardtii unravels novel complex structures. *Molecular & cellular proteomics : MCP*. **12**, 3160-83.

Mayer, M. P. (2010) Gymnastics of molecular chaperones. *Mol Cell*. **39**, 321-31.

Mayfield, S. P., A. L. Manuell, S. Chen, J. Wu, M. Tran, D. Siefker, M. Muto and J. Marin-Navarro (2007) Chlamydomonas reinhardtii chloroplasts as protein factories. *Curr Opin Biotechnol*. **18**, 126-33.

Mehrshahi, P., G. Stefano, J. M. Andaloro, F. Brandizzi, J. E. Froehlich and D. DellaPenna (2013) Transorganellar complementation redefines the biochemical continuity of endoplasmic reticulum and chloroplasts. *Proceedings of the National Academy of Sciences of the United States of America*. **110**, 12126-31.

Miao, Y. and L. Jiang (2007) Transient expression of fluorescent fusion proteins in protoplasts of suspension cultured cells. *Nature protocols*. **2**, 2348-53.

Miras, S., D. Salvi, M. Ferro, D. Grunwald, J. Garin, J. Joyard and N. Rolland (2002) Non-canonical transit peptide for import into the chloroplast. *The Journal of biological chemistry*. **277**, 47770-8.

Mitra, N., S. Sinha, T. N. Ramya and A. Surolia (2006) N-linked oligosaccharides as outfitters for glycoprotein folding, form and function. *Trends Biochem Sci*. **31**, 156-63.

Miyazono, K., J. Thyberg and C. H. Heldin (1992) Retention of the transforming growth factor-beta 1 precursor in the Golgi complex in a latent endoglycosidase H-sensitive form. *The Journal of biological chemistry*. **267**, 5668-75.

Mohorko, E., R. Glockshuber and M. Aebi (2011) Oligosaccharyltransferase: the central enzyme of N-linked protein glycosylation. *Journal of inherited metabolic disease*. **34**, 869-78.

Nada, A. and J. Soll (2004) Inner envelope protein 32 is imported into chloroplasts by a novel pathway. *Journal of cell science*. **117**, 3975-82.

Nanjo, Y., H. Oka, N. Ikarashi, K. Kaneko, A. Kitajima, T. Mitsui, F. J. Munoz, M. Rodriguez-Lopez, E. Baroja-Fernandez and J. Pozueta-Romero (2006) Rice plastidial N-glycosylated nucleotide pyrophosphatase/phosphodiesterase is transported from the ER-golgi to the chloroplast through the secretory pathway. *The Plant cell*. **18**, 2582-92.

National Science Foundation, U, USA (2013). Available at http://www.nsf.gov/pubs/2002/bio0202/model.htm

Nicol, F., I. His, A. Jauneau, S. Vernhettes, H. Canut and H. Hofte (1998) A plasma membrane-bound putative endo-1,4-beta-D-glucanase is required for normal wall assembly and cell elongation in Arabidopsis. *The EMBO journal*. **17**, 5563-76.

Okonechnikov, K., O. Golosova, M. Fursov and U. team (2012) Unipro UGENE: a unified bioinformatics toolkit. *Bioinformatics*. **28**, 1166-7.

Pagny, S., F. Bouissonnie, M. Sarkar, M. L. Follet-Gueye, A. Driouich, H. Schachter, L. Faye and V. Gomord (2003) Structural requirements for Arabidopsis beta1,2-xylosyltransferase activity and targeting to the Golgi. *The Plant journal : for cell and molecular biology*. **33**, 189-203.

Paredez, A. R., S. Persson, D. W. Ehrhardt and C. R. Somerville (2008) Genetic evidence that cellulose synthase activity influences microtubule cortical array organization. *Plant physiology*. **147**, 1723-34.

Park, M., S. J. Kim, A. Vitale and I. Hwang (2004) Identification of the protein storage vacuole and protein targeting to the vacuole in leaf cells of three plant species. *Plant physiology*. **134**, 625-39.

Parodi, A. J. (2000) Role of N-oligosaccharide endoplasmic reticulum processing reactions in glycoprotein folding and degradation. *The Biochemical journal*. **348 Pt 1**, 1-13.

Peltier, J. B., Y. Cai, Q. Sun, V. Zabrouskov, L. Giacomelli, A. Rudella, A. J. Ytterberg, H. Rutschow and K. J. van Wijk (2006) The oligomeric stromal proteome of Arabidopsis thaliana chloroplasts. *Molecular & cellular proteomics : MCP*. **5**, 114-33.

Peschke, M., D. Moog, A. Klingl, U. G. Maier and F. Hempel (2013) Evidence for glycoprotein transport into complex plastids. *Proceedings of the National Academy of Sciences of the United States of America*. **110**, 10860-5.

Powell, L. M., S. C. Wallis, R. J. Pease, Y. H. Edwards, T. J. Knott and J. Scott (1987) A novel form of tissue-specific RNA processing produces apolipoprotein-B48 in intestine. *Cell*. **50**, 831-40.

Price, J. L., D. Shental-Bechor, A. Dhar, M. J. Turner, E. T. Powers, M. Gruebele, Y. Levy and J. W. Kelly (2010) Context-dependent effects of asparagine glycosylation on Pin WW folding kinetics and thermodynamics. *Journal of the American Chemical Society*. **132**, 15359-67.

Radhamony, R. N. and S. M. Theg (2006) Evidence for an ER to Golgi to chloroplast protein transport pathway. *Trends in cell biology*. **16**, 385-7.

Raikhel, N. a. M. J. C. (2000) *Protein targeting and vesicle traffic*

Reinbothe, S., F. Quigley, A. Springer, A. Schemenewitz and C. Reinbothe (2004) The outer plastid envelope protein Oep16: role as precursor translocase in import of protochlorophyllide oxidoreductase A. *Proceedings of the National Academy of Sciences of the United States of America*. **101**, 2203-8.

Rendic, D., J. Klaudiny, U. Stemmer, J. Schmidt, K. Paschinger and I. B. Wilson (2007) Towards abolition of immunogenic structures in insect cells: characterization of a honey-bee (Apis

mellifera) multi-gene family reveals both an allergy-related core alpha1,3-fucosyltransferase and the first insect Lewis-histo-blood-group-related antigen-synthesizing enzyme. *The Biochemical journal*. **402**, 105-15.

Reumann, S., L. Babujee, C. Ma, S. Wienkoop, T. Siemsen, G. E. Antonicelli, N. Rasche, F. Luder, W. Weckwerth and O. Jahn (2007) Proteome analysis of Arabidopsis leaf peroxisomes reveals novel targeting peptides, metabolic pathways, and defense mechanisms. *The Plant cell*. **19**, 3170-93.

Rhee, S. W., T. Starr, K. Forsten-Williams and B. Storrie (2005) The steady-state distribution of glycosyltransferases between the Golgi apparatus and the endoplasmic reticulum is approximately 90:10. *Traffic*. **6**, 978-90.

Rossig, C., C. Reinbothe, J. Gray, O. Valdes, D. von Wettstein and S. Reinbothe (2013) Three proteins mediate import of transit sequence-less precursors into the inner envelope of chloroplasts in Arabidopsis thaliana. *Proceedings of the National Academy of Sciences of the United States of America*. **110**, 19962-7.

Rowland, A. A. and G. K. Voeltz (2012) Endoplasmic reticulum-mitochondria contacts: function of the junction. *Nature reviews. Molecular cell biology*. **13**, 607-25.

Ruiz-May, E., S. J. Kim, F. Brandizzi and J. K. Rose (2012) The secreted plant N-glycoproteome and associated secretory pathways. *Front Plant Sci*. **3**, 117.

Rutschow, H., A. J. Ytterberg, G. Friso, R. Nilsson and K. J. van Wijk (2008) Quantitative proteomics of a chloroplast SRP54 sorting mutant and its genetic interactions with CLPC1 in Arabidopsis. *Plant physiology*. **148**, 156-75.

Saijo, Y. (2010) ER quality control of immune receptors and regulators in plants. *Cell Microbiol*. **12**, 716-24.

Sapir-Mir, M., A. Mett, E. Belausov, S. Tal-Meshulam, A. Frydman, D. Gidoni and Y. Eyal (2008) Peroxisomal localization of Arabidopsis isopentenyl diphosphate isomerases suggests that part of the plant isoprenoid mevalonic acid pathway is compartmentalized to peroxisomes. *Plant physiology*. **148**, 1219-28.

Sarkar, M., P. A. Leventis, C. I. Silvescu, V. N. Reinhold, H. Schachter and G. L. Boulianne (2006) Null mutations in Drosophila N-acetylglucosaminyltransferase I produce defects in locomotion and a reduced life span. *The Journal of biological chemistry*. **281**, 12776-85.

Schattat, M., K. Barton, B. Baudisch, R. B. Klosgen and J. Mathur (2011) Plastid stromule branching coincides with contiguous endoplasmic reticulum dynamics. *Plant physiology*. **155**, 1667-77.

Schoberer, J., E. Liebminger, S. W. Botchway, R. Strasser and C. Hawes (2013) Time-resolved fluorescence imaging reveals differential interactions of N-glycan processing enzymes across the Golgi stack in planta. *Plant physiology*. **161**, 1737-54.

Schwacke, R., K. Fischer, B. Ketelsen, K. Krupinska and K. Krause (2007) Comparative survey of plastid and mitochondrial targeting properties of transcription factors in Arabidopsis and rice. *Molecular genetics and genomics : MGG*. **277**, 631-46.

Shaaltiel, Y., D. Bartfeld, S. Hashmueli, G. Baum, E. Brill-Almon, G. Galili, O. Dym, S. A. Boldin-Adamsky, I. Silman, J. L. Sussman, A. H. Futerman and D. Aviezer (2007) Production of glucocerebrosidase with terminal mannose glycans for enzyme replacement therapy of Gaucher's disease using a plant cell system. *Plant biotechnology journal*. **5**, 579-90.

Sheen, J (2002) A transient expression assay using *Arabidopsis* mesophyll protoplasts. Available at http://opus.mgh.harvard.edu/sheen_lab/

Sheiner, L. and B. Striepen (2013) Protein sorting in complex plastids. *Biochim Biophys Acta*. **1833**, 352-9.

Shental-Bechor, D. and Y. Levy (2008) Effect of glycosylation on protein folding: a close look at thermodynamic stabilization. *Proceedings of the National Academy of Sciences of the United States of America*. **105**, 8256-61.

Simm, S., D. G. Papasotiriou, M. Ibrahim, M. S. Leisegang, B. Muller, T. Schorge, M. Karas, O. Mirus, M. S. Sommer and E. Schleiff (2013) Defining the core proteome of the chloroplast envelope membranes. *Front Plant Sci*. **4**, 11.

Snider, (2014) Protein Glycosylation. Thermo Scientific. Available at http://www.piercenet.com /method/protein-glycosylation

Soll, J. and E. Schleiff (2004) Protein import into chloroplasts. *Nature reviews. Molecular cell biology*. **5**, 198-208.

Sparkes, I. A., J. Runions, A. Kearns and C. Hawes (2006) Rapid, transient expression of fluorescent fusion proteins in tobacco plants and generation of stably transformed plants. *Nature protocols*. **1**, 2019-25.

Spremulli, L. (2000) *Protein Synthesis, Assembly and Degradation*

Sriraman, R., M. Bardor, M. Sack, C. Vaquero, L. Faye, R. Fischer, R. Finnern and P. Lerouge (2004) Recombinant anti-hCG antibodies retained in the endoplasmic reticulum of transformed plants lack core-xylose and core-alpha(1,3)-fucose residues. *Plant biotechnology journal*. **2**, 279-87.

Stanley, P. (2011) Golgi glycosylation. *Cold Spring Harb Perspect Biol*. **3**.

Strasser, R., F. Altmann, L. Mach, J. Glössl and H. Steinkellner (2004) Generation of Arabidopsis thaliana plants with complex N-glycans lacking β1,2-linked xylose and core α1,3-linked fucose. *FEBS letters*. **561**, 132-136.

Strasser, R., J. Schoberer, C. Jin, J. Glossl, L. Mach and H. Steinkellner (2006) Molecular cloning and characterization of Arabidopsis thaliana Golgi alpha-mannosidase II, a key enzyme in the formation of complex N-glycans in plants. *The Plant journal : for cell and molecular biology*. **45**, 789-803.

Sulli, C. and S. D. Schwartzbach (1995) The polyprotein precursor to the Euglena light-harvesting chlorophyll a/b-binding protein is transported to the Golgi apparatus prior to chloroplast import and polyprotein processing. *The Journal of biological chemistry*. **270**, 13084-90.

Sung, D. Y., E. Vierling and C. L. Guy (2001) Comprehensive expression profile analysis of the Arabidopsis Hsp70 gene family. *Plant physiology*. **126**, 789-800.

Szyjanowicz, P. M., I. McKinnon, N. G. Taylor, J. Gardiner, M. C. Jarvis and S. R. Turner (2004) The irregular xylem 2 mutant is an allele of korrigan that affects the secondary cell wall of Arabidopsis thaliana. *The Plant journal : for cell and molecular biology*. **37**, 730-40.

Takahashi, J., U. J. Rudsander, M. Hedenstrom, A. Banasiak, J. Harholt, N. Amelot, P. Immerzeel, P. Ryden, S. Endo, F. M. Ibatullin, H. Brumer, E. del Campillo, E. R. Master, H. V. Scheller, B. Sundberg, T. T. Teeri and E. J. Mellerowicz (2009) KORRIGAN1 and its aspen homolog PttCel9A1 decrease cellulose crystallinity in Arabidopsis stems. *Plant & cell physiology*. **50**, 1099-115.

Tan, X., Q. Wang, B. Tian, H. Zhang, D. Lu and J. Zhou (2011) A Brassica napus lipase locates at the membrane contact sites involved in chloroplast development. *PLoS One*. **6**, e26831.

Tanz, S. K., I. Castleden, C. M. Hooper, M. Vacher, I. Small and H. A. Millar (2013) SUBA3: a database for integrating experimentation and prediction to define the SUBcellular location of proteins in Arabidopsis. *Nucleic acids research*. **41**, D1185-91.

Terasawa, K. and N. Sato (2009) Plastid localization of the PEND protein is mediated by a noncanonical transit peptide. *FEBS J*. **276**, 1709-19.

Tonkin, C. J., M. Kalanon and G. I. McFadden (2008) Protein targeting to the malaria parasite plastid. *Traffic*. **9**, 166-75.

Tonkin, C. J., N. S. Struck, K. A. Mullin, L. M. Stimmler and G. I. McFadden (2006) Evidence for Golgi-independent transport from the early secretory pathway to the plastid in malaria parasites. *Mol Microbiol*. **61**, 614-30.

Tu, B. P. and J. S. Weissman (2004) Oxidative protein folding in eukaryotes: mechanisms and consequences. *The Journal of cell biology*. **164**, 341-6.

Tu, L. and D. K. Banfield (2010) Localization of Golgi-resident glycosyltransferases. *Cellular and molecular life sciences : CMLS*. **67**, 29-41.
Tu, L., W. C. Tai, L. Chen and D. K. Banfield (2008) Signal-mediated dynamic retention of glycosyltransferases in the Golgi. *Science*. **321**, 404-7.
van Kuik, J. A., H. van Halbeek, J. P. Kamerling and J. F. Vliegenthart (1985) Primary structure of the low-molecular-weight carbohydrate chains of Helix pomatia alpha-hemocyanin. Xylose as a constituent of N-linked oligosaccharides in an animal glycoprotein. *The Journal of biological chemistry*. **260**, 13984-8.
van Wijk, K. J. and S. Baginsky (2011) Plastid proteomics in higher plants: current state and future goals. *Plant physiology*. **155**, 1578-88.
Varki, A., H. H. Freeze and P. Gagneux (2009) Evolution of Glycan Diversity. *Essentials of Glycobiology*. Varki, A., Cummings, R. D., Esko, J. D.et al (eds.) Cold Spring Harbor (NY)
Varki, A. and J. B. Lowe (2009) Biological Roles of Glycans. *Essentials of Glycobiology*. Varki, A., Cummings, R. D., Esko, J. D.et al (eds.) Cold Spring Harbor (NY)
Varki, A. and N. Sharon (2009) Historical Background and Overview. *Essentials of Glycobiology*. Varki, A., Cummings, R. D., Esko, J. D.et al (eds.) Cold Spring Harbor (NY)
Villarejo, A., S. Buren, S. Larsson, A. Dejardin, M. Monne, C. Rudhe, J. Karlsson, S. Jansson, P. Lerouge, N. Rolland, G. von Heijne, M. Grebe, L. Bako and G. Samuelsson (2005) Evidence for a protein transported through the secretory pathway en route to the higher plant chloroplast. *Nat Cell Biol*. **7**, 1224-31.
Voinnet, O., S. Rivas, P. Mestre and D. Baulcombe (2003) An enhanced transient expression system in plants based on suppression of gene silencing by the p19 protein of tomato bushy stunt virus. *The Plant journal : for cell and molecular biology*. **33**, 949-56.
von Schaewen, A., J. Frank and H. Koiwa (2008) Role of complex N-glycans in plant stress tolerance. *Plant signaling & behavior*. **3**, 871-3.
von Schaewen, A., A. Sturm, J. O'Neill and M. J. Chrispeels (1993) Isolation of a mutant Arabidopsis plant that lacks N-acetyl glucosaminyl transferase I and is unable to synthesize Golgi-modified complex N-linked glycans. *Plant physiology*. **102**, 1109-18.
Waheed, A., T. Okuyama, T. Heyduk and W. S. Sly (1996) Carbonic anhydrase IV: purification of a secretory form of the recombinant human enzyme and identification of the positions and importance of its disulfide bonds. *Archives of biochemistry and biophysics*. **333**, 432-8.
Waller, R. F., M. B. Reed, A. F. Cowman and G. I. McFadden (2000) Protein trafficking to the plastid of Plasmodium falciparum is via the secretory pathway. *The EMBO journal*. **19**, 1794-802.
Wang, Y., J. Tan, M. Sutton-Smith, D. Ditto, M. Panico, R. M. Campbell, N. M. Varki, J. M. Long, J. Jaeken, S. R. Levinson, A. Wynshaw-Boris, H. R. Morris, D. Le, A. Dell, H. Schachter and J. D. Marth (2001) Modeling human congenital disorder of glycosylation type IIa in the mouse: conservation of asparagine-linked glycan-dependent functions in mammalian physiology and insights into disease pathogenesis. *Glycobiology*. **11**, 1051-70.
Wang, Z. and C. Benning (2012) Chloroplast lipid synthesis and lipid trafficking through ER-plastid membrane contact sites. *Biochem Soc Trans*. **40**, 457-63.
Wilkinson, D. A. (1998) Getting the message with RT-PCR. *The Scientist*. **12**, 20–22.
Wilson, I. B. (2001) Identification of a cDNA encoding a plant Lewis-type alpha1,4-fucosyltransferase. *Glycoconjugate journal*. **18**, 439-47.
Wisniewski, J. R., A. Zougman, N. Nagaraj and M. Mann (2009) Universal sample preparation method for proteome analysis. *Nature methods*. **6**, 359-62.
Wong, K. F. and J. Ho (2013) Stabilization of ferrochelatase via lysine residues on the carboxyl terminal extension. *Protein and peptide letters*. **20**, 977-81.

Wujek, P., E. Kida, M. Walus, K. E. Wisniewski and A. A. Golabek (2004) N-glycosylation is crucial for folding, trafficking, and stability of human tripeptidyl-peptidase I. *The Journal of biological chemistry*. **279**, 12827-39.

Xu, C., J. Fan, A. J. Cornish and C. Benning (2008) Lipid trafficking between the endoplasmic reticulum and the plastid in Arabidopsis requires the extraplastidic TGD4 protein. *The Plant cell*. **20**, 2190-204.

Yang, X., L. Tu, L. Zhu, L. Fu, L. Min and X. Zhang (2008) Expression profile analysis of genes involved in cell wall regeneration during protoplast culture in cotton by suppression subtractive hybridization and macroarray. *Journal of experimental botany*. **59**, 3661-74.

Zhao, X., G. Li and S. Liang (2013) Several Affinity Tags Commonly Used in Chromatographic Purification. *J Anal Methods Chem*. **2013**, 581093.

Zhou, C., J. G. Tokuhisa, D. R. Bevan and A. Esen (2012) Properties of beta-thioglucoside hydrolases (TGG1 and TGG2) from leaves of Arabidopsis thaliana. *Plant Sci*. **191-192**, 82-92.

Zielinska, D. F., F. Gnad, K. Schropp, J. R. Wisniewski and M. Mann (2012) Mapping N-glycosylation sites across seven evolutionarily distant species reveals a divergent substrate proteome despite a common core machinery. *Mol Cell*. **46**, 542-8.

Zielinska, D. F., F. Gnad, Wi and M. Mann (2010) Precision Mapping of an In Vivo N-Glycoproteome Reveals Rigid Topological and Sequence Constraints. *Cell*. **141**, 897-907.

Zouhar, J., E. Rojo and D. C. Bassham (2009) AtVPS45 is a positive regulator of the SYP41/SYP61/VTI12 SNARE complex involved in trafficking of vacuolar cargo. *Plant physiology*. **149**, 1668-78.

Zybailov, B., H. Rutschow, G. Friso, A. Rudella, O. Emanuelsson, Q. Sun and K. J. van Wijk (2008) Sorting signals, N-terminal modifications and abundance of the chloroplast proteome. *PLoS One*. **3**, e1994.

SUPPLEMENTAL FIGURES: CHAPTER 1

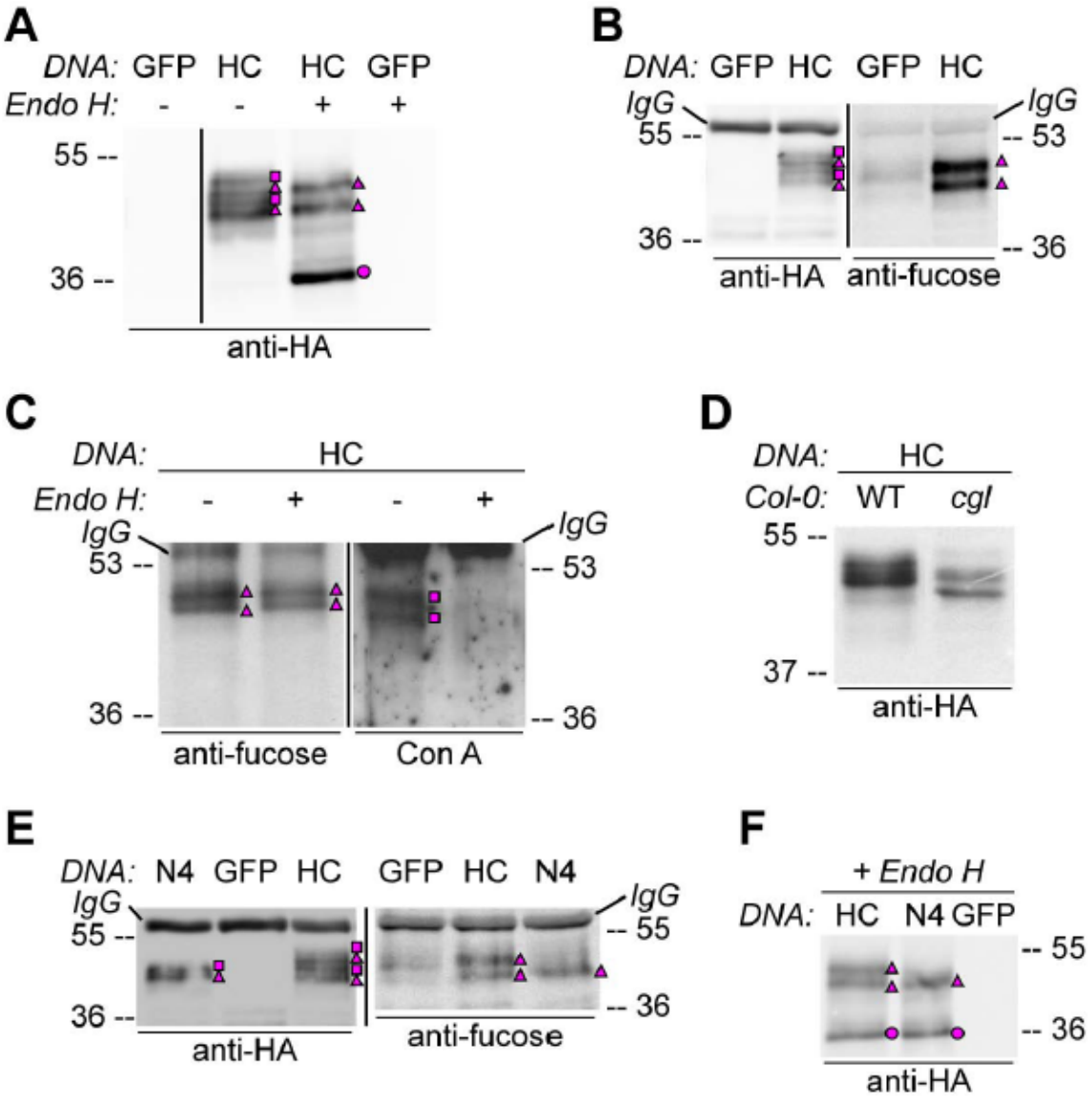

Figure S1.1. HA-tagged CAH1 contains both high mannose-type and complex-type N-glycans. Analysis of protein extracts of protoplasts from Arabidopsis suspension culture cells transiently expressing GFP (negative control), HA-tagged wt (HC) or single mutated (N4) CAH1. Heavy chain from immunoprecipitation is indicated as IgG. (A) HC protein migrates as four distinct bands: two isoforms are sensitive (squares) and two resistant (triangles) to Endo H. Deglycosylated protein is labelled with a dot. (B) HA-immunoprecipitation of HC reveals four protein isoforms when probed with HA antibodies, (squares and triangles) two of which are recognized by a(1,3)-fucose antibodies (triangles). (C) Endo H treatment of HC shows that the fucose containing glycoforms are resistant to the enzyme (triangles). The glycoforms detected by Con A affinoblot (squares) are sensitive to Endo H, corresponding to high mannose-type glycoforms. (D) HC migrates as two main bands when expressed in Columbia (Col)-0 mesophyll protoplasts from the cgl mutant, which is unable to produce complex-type N-glycans. (E) HA immunoprecipitation of HC reveals four protein isoforms, while only two can be detected of the N4 mutant using HA antibodies (triangles and squares). Analysis using a(1,3)-fucose antibodies on the same samples showed that only two of the HC, and one of the N4, glycoforms harbour complex-type glycans (triangles). (F) Endo H treatment and HA immunodetection of total protein extract confirmed that only one glycoform in N4 and two in HC are resistant to the enzyme (triangles). Deglycosylated isoforms are marked with dots. From Burén *et al.*, 2001

SUPPLEMENTAL FIGURES: CHAPTER 3

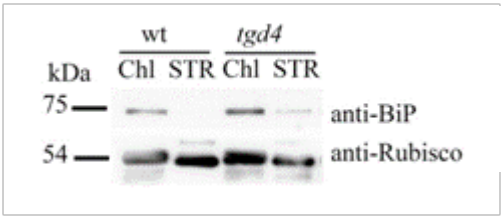

Figure S3.1: Stroma fraction purification/enrichment. Immnoblots of stroma fractions of wild type and tgd4 mutant using anti-BiP (ER marker) and anti-Rubisco as loading control. 10 µg of protein per lane. wt, wild type, Chl, chloroplasts, STR, stroma.

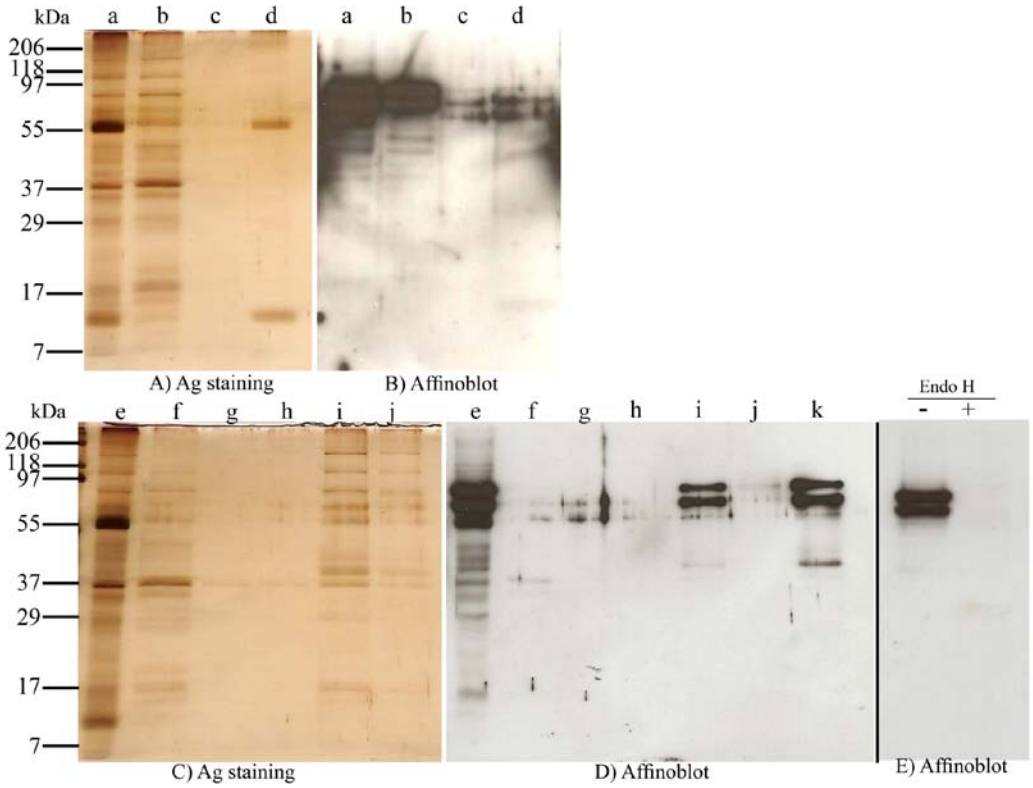

Figure S3.2: Purification of wild type stroma samples after SepproRubisco Spin Colums (A and B) and Con A affinity columns fractionation (C and D). A) Silver staining (Ag), 1 µg of protein per lane. **B)** Affinoblot, 2 µg of protein per lane. a) Initial stroma from wt; b) Flow through containing most proteins from the original sample except Rubisco, immobilized in the column; c) Wash fraction containing proteins non-specifically bound to column; d) Eluted fraction containing proteins specifically bound to column, mostly Rubisco. **C)** Silver staining (Ag), 1 µg of protein per lane, **D)**Affinoblot, 2 µg of protein per lane. e) Initial stroma fraction, f) Flow-through containing non-glycosylated proteins, Washing steps 1(g)and2 (h) containing proteins non-specifically bound to the column; Eluted fractions1 (i) and 2(j) containing N-glycoproteins. Is in this fraction where most of HMGP are placed. k) Total N-glycoproteins pooled from i and j fractions, further subjected to MALDI-TOF protein identification. **E)** Endo H treatment treated (+) and non-treated (-) Con A columns eluted fraction (k) verifying the presence of high mannose N-glycoproteins.

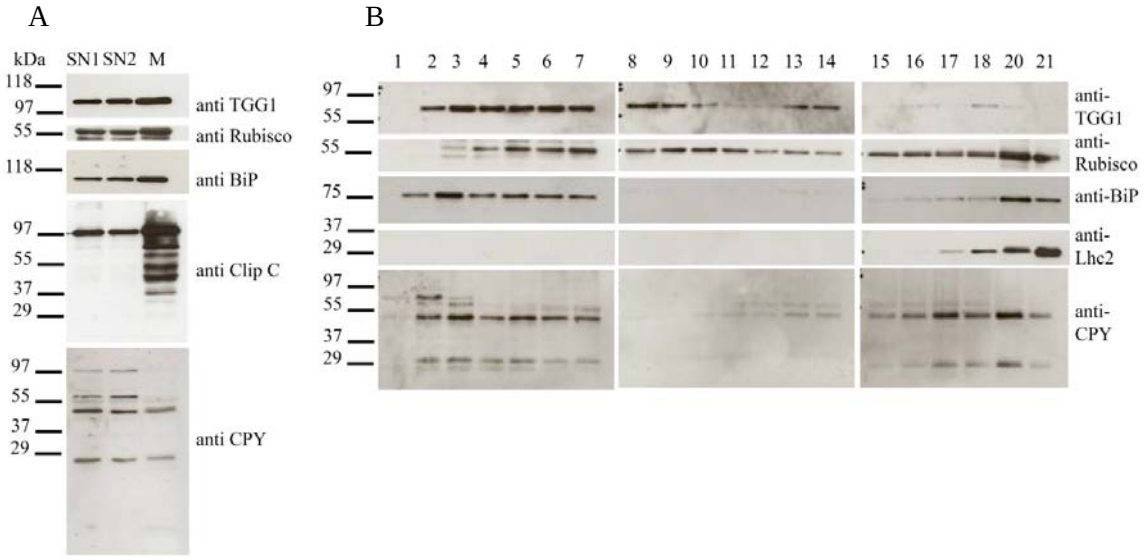

Figure S3.3: Enrichment of TGG1 in microsome fractions and distribution in sucrose gradients A) Analysis of subcellular fractions using anti-TGG antibodies and different organelle markers (anti-Rubisco, chloroplast stroma; anti-BiP, ER; anti-Clip C, chloroplast envelope; anti-CPY, vacuole). B) Microsome gradients showing distribution of TGG1 and organelle markers (Lhc2, thylakoid specific marker). SN1, cell total extract, SN2, soluble cellular fraction, M, micrososme fraction

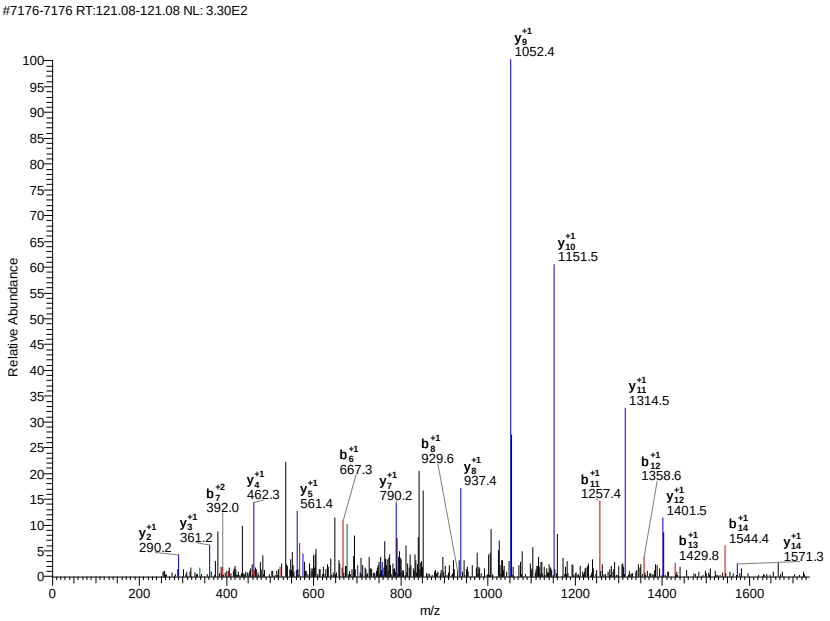

Figure S3.4: Mass spectrum of an identified peptide from TGG2. Graphic shows the m/z ratio of charged fragments (ions) from a peptide which has been fragmented in their peptide bonds. Differences in mass in successive peaks allow identification of the sequence of the peptide. b-type ions are fragments whose charge is retained in amino terminal from the peptide bond, while y-type ions are those retained in carboxyl terminus of the peptide bond.